1963

Statistical procedures applicable in the analysis of bioassays when the usual assumptions are not fulfilled

Paul E. Leaverton
Iowa State University

STATISTICAL PROCEDURES APPLICABLE IN THE ANALYSIS OF BIOASSAYS WHEN THE USUAL ASSUMPTIONS ARE NOT FULFILLED

by

Paul Emmett Leaverton

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of
The Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Major Subject: Statistics

Approved:

Signature was redacted for privacy.
In Charge of Major Work

Signature was redacted for privacy.
Head of Major Department

Signature was redacted for privacy.
Dean of Graduate College

Iowa State University
Of Science and Technology
Ames, Iowa

1963

TABLE OF CONTENTS

I. INTRODUCTION

A. Statistical Analysis of Bioassays

A generally accepted definition of the term bioassay (biological assay) as it is used in statistics is that given by Finney (15, p. 1). He states that, "Biological assays are methods for the estimation of the nature, constitution, or potency of a material (or of a process) by means of the reaction that follows its application to living matter . . . Quantitative assays, with which we are concerned, are similar to methods of physical measurement or of quantitative chemical analysis in that their function is to provide numerical assessment of some property of the material to be assayed. An essential part in this assessment is played by measurement of growth or other changes in animals, plants, animal tissue, micro-organisms, or some other form of living matter. An assay is thus a form of biological experiment, but the interest lies in comparing the magnitudes of effects of different treatments. The experimental technique may be the same as is used in a purely comparative experiment but the difference in purpose will affect the optimal experimental design and the statistical analysis." For example, an investigation which simply aims at determining whether or not two toxins are different is not necessarily a bioassay by the above definition. It becomes one if the experimenter is additionally

interested in estimating the relative potency of the toxins.

Bioassays may be classified as being either direct or indirect. The principle of a direct assay is that the dose of stimulus just sufficient to produce a specified response is measured directly. This type of assay is not common because technical problems often prevent the experimenter from ensuring that the subjects receive the exact dose required.

In the more common indirect assays, different doses are administered, each to several subjects and the resulting responses are observed. If an individual subject's response is an "all or none" reaction (such as survival or death) it is called a quantal response. Quantal response assays are distinguished from quantitative response assays in which the magnitude of some property of the subject (such as weight) is recorded. This thesis will only be concerned with quantal and quantitative response indirect assays.

Bioassays have found application in a great many scientific fields. Their use can be noted with increasing frequency in the literature of immunology, pharmacology, endocrinology, entomology, and many other areas of biological and medical research. The technique of comparing the effect of an experimental preparation with that of a standard preparation has proven to be of great value in helping to "remove" much of the variation between experimental units from the comparison of two preparations. In bioassays, a specially calculated comparison, usually called the relative potency, is expressed

as a ratio of equally effective doses. It is sometimes known as relative activity or by some other term depending upon the field of application.

The statistical analysis of a bioassay involves the following:

1. specification of the mathematical model and its related assumptions
2. testing, in some sense, the validity of the model and assumptions
3. providing an estimate of relative potency along with appropriate confidence limits.

The success of a bioassay in providing a numerical assessment of a drug or other material depends to a great extent upon the relevance of the model employed to the real situation. Frequently a convenient linear model is used as an approximation to the dose-response relationship. Transformations, sometimes called metameters, of both the dose and response variables are often used to achieve this end.

A considerable amount of research on the development of statistical methods of analysis pertinent to bioassays has been and is being done. An excellent exposition of the basic concepts used in bioassays was published by Jerne and Wood (21) in 1949. In this paper they enumerate and examine in detail a number of the implicit and explicit assumptions involved in using a bioassay model. A standard text in this field, which consolidated most of the work through 1952, is

that by Finney, "Statistical Method in Biological Assay" (15). In this comprehensive book the prominent mathematical models and methods in use are discussed and criticized. Cox (11) lists over one hundred and fifty references dealing with statistical methods in bioassay, most of which have been published since 1950. A large number of even more recent developments are available in the somewhat scattered literature and are not cited here. References to works which are pertinent to this study are made later where appropriate.

B. The Tolerance Distribution in Quantal Response Bioassays

The concept of a tolerance distribution has been found useful in establishing mathematical models for use in quantal response assays. In theory there is an exact dose which is just sufficient to cause the response in an individual subject. This dose, called the tolerance, varies between subjects even within a homogeneous population of subjects. If these tolerances could be directly measured, as in a direct assay, they would have a frequency distribution. The fact that such direct measurements would usually be quite difficult to obtain has already been pointed out. Ordinarily several different doses are administered, each to a group of subjects which are randomly selected from a population. The percentage which has responded at each dose is then observed. This provides an estimate of the percentage of the parent popula-

tion which have tolerances less than or equal to that particular dose. The series of percentages thus obtained approximates the cumulative distribution function of the frequency distribution of tolerances.

Reference is often made to an LD-50 (lethal dose-50), ED-50 (effective dose-50) or a fifty per cent end point in quantal response studies. These all refer to the median, equivalent to the mean in symmetrical distributions, of the hypothetical tolerance distribution. Several different methods have been developed for the estimation of this quantity. Some specify a parametric model for the cumulative tolerance distribution such as the cumulative normal or the logistic distribution. In such cases, if necessary to facilitate estimation, a transformation, such as a probit or logit, is made to give a response metameter having a linear relationship with a dose metameter (e.g. log dose).

In the literature, a number of estimators of the LD-50 have been proposed which do not specify the exact form of the tolerance distribution but can be applied to broad classes of distributions, such as all symmetric distributions. Whether a specific distributional form is used as a model or one of these "distribution-free" estimators depends upon several factors. Probably the major determining factors in this regard are available knowledge of the distributional form and the computational convenience of the methods. Finney (14) has reviewed the advantages and disadvantages of most of the

various methods.

C. Some Limitations of Presently Available Bioassay Models and Techniques

1. Monotonicity

The general situation in a bioassay may be described as follows. Let y_i denote the response of a subject, randomly selected from a population, receiving a known dose z_i of a preparation. Let also $E(y_i | z_i) = Y_i$. The dose-response relationship may be expressed as $Y_i = F(z_i)$, or $Y = F(z)$, where $F(z)$ represents a function of z. The assumption that $F(z)$ is a strictly monotonic function of z in the usual dose range is a basic requirement of all standard bioassay models. That is, for every z_i, z_j, if $z_i < z_j$ then $F(z_i) < F(z_j)$. It is necessary that this assumption be true in order to insure that comparisons of the effects of different doses (or treatments), even at a specific level of Y, be unique and meaningful. The occurrence of non-monotonic dose-response functions in studies of the effects of Cerium Edetate on mice and guinea pigs has been noted by Dr. Joseph Graca of Ames, Iowa, in 1963 (private communication).

2. Similarity

Another assumption necessary for the estimation of relative potency by the usual bioassay techniques is known as the condition of similarity. For this condition to be true, the test preparation (T) must behave as though it were a dilution

(or a concentration) of the standard preparation (S) in a diluent that is completely 'inert' with respect to the response used. In this case, it follows that the relative potency, $\rho = \frac{z_S}{z_T}$ (the ratio of equally effective doses) is constant for all Y. The algebraic statement of this condition as it affects the dose-response curves is:

$$F_T(z_T) = F_S(z_S) = F_S(\rho z_T) \qquad (1)$$

for all z_S, z_T.

If, in this situation, the response curve for the test preparation T is $Y = \alpha + \beta_T z_T$, where Y is some response metameter, then by (1), the curve for the standard preparation S is $Y = \alpha + \beta_S z_S = \alpha + \beta_S(\rho z_T)$. This is the standard slope ratio assay model where ρ is estimated by the ratio of the slopes of the two fitted regression lines. The intercept, α, is common to both lines since it is the expected response to a zero dose in each case.

If a logarithmic transformation of the dose is made, where $x = \log_{10}$ dose (z), then $z = 10^x$ and, under the condition of similarity, it can be seen that

$$F_T(10^{x_T}) = F_S(\rho 10^{x_S}) = F_S(10^{x_T} + \log \rho),$$

as noted by Cox (11). This shows that whatever the form of the response curve for T, the curve for S is the same except for a horizontal displacement by an amount $\log \rho$. When, more specifically, a linear model is appropriate, parallel regression lines are usually fitted to the two sets of observations

and an estimate of log ρ is taken as the horizontal distance between these two lines. A bioassay in which such a model is employed is called a parallel line assay.

The requirement of similarity for a "valid" assay has been stressed by Finney, as quoted by Jerne and Wood (21, p. 277), "If the data cannot be adequately described by the same form of F for both preparations, the basic assumption that only the same effective constituents were concerned in both must be false. The assay is therefore invalid." Jerne and Wood (21, p. 277) comment, ". . . and it might be added that the whole idea of assaying that particular test preparation against that particular standard preparation becomes absurd." Although it should be obvious that validity in this sense means validity of the bioassay model to provide a single value for relative potency, some confusion on this point is apparent. Biologists sometimes find themselves at a loss in attempting to interpret bioassay data where the condition of similarity clearly does not apply. For example, Gibbs et al. (16, p. 408), having encountered such a situation, go so far as to state that, ". . . no quantitative comparisons can be made . . ."

Of course, as most experimenters realize, nature is under no obligation to provide data that fit our mathematical models. Frequently these are only models of computational and interpretive convenience. There may be no theoretical justification for believing that two chemically different substances behave

as though one were a diluent of the other. Yet nearly all of the statistical literature available on bioassays accept this notion as a first requisite. The only exception known to the author is the paper by Thompson (29). Some biologists have recognized this limitation of bioassay models, and noting such "invalidity" in their own work, have asked for better models. This situation is clearly illustrated by Grimshaw and D'Arcy (18, p. 262). Having observed nonparallel log dose-response lines (in violation of the condition of similarity) in their pharmacological studies, they suggest the possibility of a new index that takes this fact into account. They conclude, "It would therefore seem that whilst there are several methods available to detect local anaesthetic activity there is not as yet a fully satisfactory method for a complete quantitative assessment."

3. Linearity

In the majority of bioassays, it is assumed that the relationship $Y = F(z)$ is linear or may be made linear by appropriate transformations of Y and z. A major difficulty, as one might suspect, in some exploratory bioassays is that of finding transformations (metameters) of the dose and response units so that a linear model is appropriate. If time and other resources were always available to the extent that the form of the dose-response curve could be well established, there would be few problems in this regard. However, situa-

tions exist where such an examination is not possible. In some studies, for example, large quantities of test preparations are screened. The nature of the dose-resonpse curves are unknown and only a few dose levels for each preparation may be used. The use of any specific parametric models for $F(z)$ in such cases may be suspect.

Finney, (15, p. 162) in 1952, and Bliss, (5) in 1957, discuss the use of a quadratic model in cases where a linear model appears to be inadequate. Mantel and Hilgar (24, p. 61) have considered the effects of using a linear model in bioassays when, in fact, $F(z)$ is nonlinear. Their study is concerned with the routine screening of drugs with bioassay techniques. They point out the fact that, "The existence of such nonlinearity could lead to extremely misleading estimates of relative potency if straight line procedures were routinely used with all data received." In the absence of a parametric model, their empirical study resulted in a rule of thumb procedure that appeared to work fairly well with a specific sigmoidal curve.

When situations arise where there is little or no theoretical basis for a specific dose-response model, one can well establish a case for estimators which do not require such a model. Such estimators are frequently used in quantal response assays. However, as Finney (14) points out, all of the standard methods of this type have serious shortcomings. For example, some lack procedures for testing the validity of

the assumptions, and some require unrealistic assumptions. A few specific examples will be discussed in Chapter III. Recent studies in this area are by Van Eeden (32), Goldberg (17), and Cochran (10).

The limitations of a mathematical model are determined by the validity of the assumptions inherent in the model. Three basic assumptions in the standard bioassay models that have been pointed out in this chapter are:

1. monotonicity of $F(z)$
2. the condition of similarity
3. linearity of $Y = F(z)$, where Y is the response metameter and z is the dose metameter.

Assumptions 1 and 2 are prerequisites in all of the standard models. Assumption 3 is required of most assay models such as in slope ratio or parallel line assays.

A few examples of situations occurring in the practical application of bioassays where the above assumptions failed to hold have been cited. The purpose of this thesis has been to provide statistical procedures relevant to the consideration of these basic assumptions. In Chapter II tests for monotinicity and similarity, which are not based on a parametric form of $F(z)$, are described. Chapter III is concerned with the development of a distribution-free estimator of relative potency in quantal response assays. The derivation of a

general model which offers a means of expressing the results of bioassays where the condition of similarity does not apply is given in Chapter IV.

II. BASIC VALIDITY TESTS

A. Tests for Monotonicity

Validity tests, in the statistical sense, are significance tests of the hypothesis that the observations are compatible with the assumptions required by the chosen mathematical model. Tests of the hypothesis that $F(z)$ is monotonic, where the exact parametric form of $F(z)$ is unspecified, are given in this chapter. Ideally, it would be preferable to test for strict monotinicity, however, the general manner in which the null hypothesis must be specified forces consideration of tests for monotonicity only. The null hypothesis of monotonicity will be accepted unless significant evidence against H_o is observed at the dose levels employed. It should, however, be remembered, that no information on $F(z)$ between the actual z's used is available.

Van Eeden (31), in 1960, considered a class of tests of the hypothesis that k parameters

$$\theta_1, \theta_2, \ldots, \theta_k$$

satisfy the inequalities

$$\theta_1 \leq \theta_2 \leq \ldots \leq \theta_k$$

against the alternative hypothesis H_A that at least one value of i exists for which $\theta_i > \theta_{i+1}$. Although this null hypothesis H_o is appropriate for testing the monotonicity of $F(z)$, the alternative hypothesis here is not and leads to difficulties

as will be shown.

The situation to be considered may be stated as follows. At each known independent dose z_i, there are n_i independent observations ($i = 1, \ldots, k$). Let y_{ij} ($j = 1, \ldots, n_i$) represent an observation at z_i and let θ_i denote the mean of the distribution from which y_{ij} is drawn. If at each i, this distribution is normal with a known variance σ^2, Van Eeden's test criterion leads to use of the statistic $d_i = \bar{y}_{i+1} - \bar{y}_i$ ($i = 1, \ldots, k - 1$), where $\bar{y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij}$. A critical region is determined for each d_i and H_o is rejected if any d_i exceeds its critical value. No account is taken of the magnitude of k.

Because of the form of the alternative hypothesis, this test obviously has low power against the particular alternative of a gentle downward trend among the θ_i's. Specifically, it may be seen that H_o will not be rejected by this test criterion even though $\bar{y}_1 > \bar{y}_2 > \ldots > \bar{y}_k$, such that $\bar{y}_1 > > \bar{y}_k$ provided no d_i is significant. Although this test procedure may be appropriate for the given H_A, it appears, therefore, that it is not appropriate for the general bioassay problem as Van Eeden (32, p. 209) has implied.

A related test procedure was given by Bartholomew (3), in 1959. In this case the hypothesis to be tested is $H_o : \theta_1 = \theta_2 = \ldots = \theta_k$ and the alternative hypothesis is $H_A : \theta_1 \leq \theta_2 \leq \ldots \leq \theta_k$ where at least one of the inequality

signs must be a strict inequality. In subsequent papers (3, 4), Bartholomew has discussed some properties of his derived test criterion. Although, this is not the hypothesis of interest here, it has analogies with one developed in Section C which is specifically adapted to bioassay.

B. The Quantal Response Case

1. A range test for monotonicity

The experimental situation may be described as follows. Let the number of independent observations n, be the same at each independent dose z_i ($i = 1, \ldots, k$). The y_{ij} ($j = 1, \ldots, n$) observations at each z_i are assumed to be binomially distributed with parameters P_i and n. Let $y_{ij} = 1$ if the response occurred and $y_{ij} = 0$ if the response did not occur. The P_i's are assumed to lie on the dose-response curve $P_i = F(z_i)$, where $F(z_i)$ is a continuous non-decreasing function (the non-increasing case can be treated by making certain obvious sign changes). No parametric form for the relationship $P_i = F(z_i)$ is specified.

In order to test the assumption that $F(z_i)$ is non-decreasing, the following hypothesis $H_o : P_i \leq P_2 \leq \ldots \leq P_k$ is tested against the alternative hypothesis, $H_A : P_i > P_{i+m}$ for at least one i and m ($i = 1, \ldots k$; $m = 1, \ldots k - i$). An intuitively reasonable statistic for use in such a test is observed value of $\max_{\text{all } i,\, m} (\hat{P}_i - \hat{P}_{i+m})$, where the symbol $\hat{}$ over

a parameter indicates an estimate of it. Since n is the same at each z_i, an equivalent statistic is $\max_{\text{all } i,\, m} (Y_i - Y_{i+m})$ where $Y_i = \sum_{j=1}^{n} y_{ij}$, $Y_{i+m} = \sum_{j=1}^{n} y_{i+m,\, j}$. If this quantity is not positive it gives no evidence contrary to H_o and H_o would be accepted. In cases where the quantity is positive it then becomes necessary to examine its distribution under H_o. It can be seen that the use of $\max_{\text{all } i,\, m} (Y_i - Y_{i+m})$ as a test statistic leads to a one tailed test, since differences in one direction only are of interest.

Consider the distribution of $w = Y_{max} - Y_{min}$ under the special case of the null hypothesis, H_o' : $P_1 = P_2 = \ldots = P_k = 0.5$. A general expression for the probability that the range w equals r, when k independent samples each of size n are drawn from a binomial population with probability of a response equal to P, is given by

$$P_r(w=r) = \sum_{i=o}^{n-r} \left\{ \left[\sum_{j=i}^{i+r} \binom{n}{j} P^j Q^{n-j} \right]^k - \left[\sum_{j=i+1}^{i+r} \binom{n}{j} P^j Q^{n-j} \right]^k - \left[\sum_{j=i}^{i+r-1} \binom{n}{j} P^j Q^{n-j} \right]^k + \left[\sum_{j=i+1}^{i+r-1} \binom{n}{j} P^j Q^{n-j} \right]^k \right\}. \quad (2)$$

The above expression can be derived by considering the probability that all of the observations fall into the interval [i, i+r] where the values of i and i+r are obtained at least once.

The probability, for a given i, that all of the observations are in [i, i+r] and that values i and i+r are obtained at least once, is Pr {all observations are in [i, i+r]} - Pr {all observations are in [i+1, i+r] or [i, i+r - 1]}. These probabilities are then summed over all i.

The special case of (2) of interest here is obtained by putting P = 0.5. In this case

$$P(w=r) = \left(\frac{1}{2}\right)^{nk} \sum_{i=0}^{k-r} \left\{ \left[\sum_{j=1}^{i+r} \binom{n}{j} \right]^k - \left[\sum_{j=i+1}^{i+r} \binom{n}{j} \right]^k - \left[\sum_{j=i}^{i+r-1} \binom{n}{j} \right]^k + \left[\sum_{j=i+1}^{i+r-1} \binom{n}{j} \right]^k \right\}. \qquad (3)$$

The use of (3) is illustrated in the following example. Given n = 4, k = 4, then

$$Pr(w=2) = \left(\frac{1}{2}\right)^{16} \sum_{i=0}^{2} \left\{ \left[\sum_{j=i}^{i+2} \binom{4}{j} \right]^4 - \left[\sum_{j=i+1}^{i+2} \binom{4}{j} \right]^4 - \left[\sum_{j=i}^{i+1} \binom{4}{j} \right]^4 + \left[\sum_{j=i+1}^{i+1} \binom{4}{j} \right]^4 \right\}$$

$$= \left(\frac{1}{2}\right)^{16} [11^4 - 10^4 - 5^4 + 4^4 + 14^4 - 10^4 - 10^4 + 6^4 + 11^4 - 10^4 - 5^4 + 4^4]$$

$$= \left(\frac{1}{2}\right)^{16} (28, 256)$$

$$= 0.431.$$

In theory, the exact distribution of w could be calculated for restricted combinations of n and k likely to be used in certain types of bioassays, and it would not be a formidable task for an electronic computer to obtain the probabilities.

As an alternative to obtaining these exact probabilities, the normal approximation to the binomial, known to be quite good when P = 0.5 even in small samples, was investigated. Pearson (26) has tabulated some percentage points for the distribution of the standardized range $W = \frac{Y_{max} - Y_{min}}{\sigma}$ in samples of size k (k = 1, 2, . . ., 20) from a normal population with a known σ. Application of the normal approximation yields the following results:

$$\frac{Y_{max} - Y_{min}}{\sqrt{nPQ}} = \frac{Y_{max} - Y_{min}}{\sqrt{(n)(.5)(.5)}} = W_k$$

$$Y_{max} - Y_{min} = \frac{W_k}{2}\sqrt{n}$$

$$Y_{max} - Y_{min} = C_k \sqrt{n}$$

It follows, therefore, that in order to obtain the 5 per cent and 1 per cent critical points, C_{k_5} and C_{k_1} respectively, the tabulated W_k's must be halved. Since the test of interest is a one tailed test, the appropriate W_k's are those corresponding to percentage points 10 and 2. Table 1 presents values of C_{k_α} obtained in this manner. A test of H_o^* against the alternative H_A is thus provided by comparing max $(Y_i - Y_{i+m})$ observed in k samples of size n with

$C_{k_\alpha}\sqrt{n}$.

A comparison of the upper percentage points of the exact and approximate distribution of w was made for two cases. In one case where k = 4 and n = 4, the exact probability that w = 4 is .04, while the probability for this event provided by the normal approximation is .06. In the case where k = 10 and n = 4, $Pr(w \geq 7) = .036$ is the exact figure and the approximate probability is .021. The approximation thus appears to be reasonable even for such small samples as n = 4. Further preliminary numerical investigation suggests that the tail of the exact distribution of w is somewhat thicker than that of the corresponding approximate distribution. However, from the closeness of the approximation, in cases so far examined, it appears that use of the test based on the normal approximation would be reasonable in most practical cases. As an additional safeguard against obtaining misleading inferences, it may be noted that the actual null hypothesis here tested is itself conservative in the sense that the probability of rejecting H_o, and accepting H_A, is greater in the case H_o' than for any other case under H_o. This can be seen by noting that in a case where an equality sign in H_o' is replaced by an inequality of the type that occurs in H_o, then

$$Pr\left[\max_{\text{all } i,\, m} (Y_i - Y_{i+m}) > C\right]$$

for any real constant C, would not be greater than the probability of this event under H_o'. Critical valves established

Table 1. Values of C_{k_α} (α = .05, .01) for the approximate one tailed range test in binomial populations

(k) Number of samples	C_{k} .05	C_{k} .01	(k) Number of samples	C_{k} .05	C_{k} .01
2	1.160	1.645	11	2.105	2.475
3	1.450	1.900	12	2.140	2.510
4	1.620	2.045	13	2.175	2.540
5	1.735	2.150	14	2.205	2.565
6	1.830	2.230	15	2.235	2.590
7	1.900	2.295	16	2.260	2.615
8	1.965	2.350	17	2.285	2.635
9	2.020	2.400	18	2.305	2.655
10	2.065	2.440	20	2.345	2.695

for P = 0.5 will also, of course, be conservative in the above sense if the true value of P is not equal to 0.5.

An example illustrating the range test uses data taken from a study by Carlson (9) in 1963. In an investigation of the effects of an insecticide (Co - Ral) on a certain species of insects (Hexagenia), the following results were observed, where n = 30 at each dose.

Dose (coded)	Number dead
1	$Y_1 = 4$
2	$Y_2 = 12$
3	$Y_3 = 17$
4	$Y_4 = 21$
5	$Y_5 = 19$
6	$Y_6 = 23$
7	$Y_7 = 30$

The $\max_{\text{all } i,\, m} (Y_i - Y_{i+m})$ is $21 - 19 = 2$, and $C_{7,\ .05}\sqrt{n}$ $= 1.90\sqrt{30} = 10.4$. The hypothesis that $P_1 \leq P_2 \leq \ldots \leq P_7$ is accordingly accepted, in this case.

It might be remarked in passing that the methods described in this chapter for determining significance levels of the range in samples from a binomial population may have applications in making "multiple range tests" in other types of analyses based on proportions. In fact, the idea of determining which $\hat{P}_i$'s are significantly different, by comparing $\hat{P}_i$ and $\hat{P}_j$ (for all i, j) with the percentage points for the distribution of the range, is directly analogous to Tukey's (30) multiple range test.

2. An alternative test for monotonicity

The range test for monotonicity is obviously of low power with respect to certain specific alternatives under H_A. This is a result of the generality of the hypotheses and the fact that the "worst" possible case under H_o had to be considered.

Another criticism of the range test is that the test statistic is based only upon the maximum discrepant observation, and does not utilize all of the evidence against H_o unless $Y_i > Y_{i+1}$ for only one i. If $Y_i > Y_{i+1}$ for more than one i, it is desirable to have a test which uses more of the information available. An appropriate procedure based initially on that suggested by Bartholomew (3), is derived in this section, and although it is not clear whether or not the test has optimum properties, the technique appears to give meaningful inferences.

Suppose, as before, that results of k independent binomial samples, of n trials each, are given, where the number having responded at trial z_i is Y_i, the hypothesis

$$H_o: P_1 \leq P_2 \leq \ldots \leq P_k$$

is to be tested against

$$H_A: P_i < P_{i+m}$$

for at least one i and m (i = 1, . . ., k) (m = 1, . . ., k-i).

The maximum likelihood estimates p_i' of the P_i, under H_o, are the same as the unrestricted maximum likelihood estimates of P_i, namely $\frac{Y_i}{n}$, if $Y_1 < Y_2 < \ldots < Y_k$. If, however, for value(s) of i, $Y_i > Y_{i+1}$, then in order to obtain the p_i''s, the i+h and i+1+h samples are averaged giving $p_i' = p_{i+1}' = \frac{Y_i + Y_{i+1}}{2n}$. Starting with Y_k and progressing to Y_1, this procedure is continued, an averaged p' being treated as a single observation (weighted by its total n), until estimates which fulfill the condition:

$$p_1' \leq p_2' \leq \cdots \leq p_k'$$

are obtained. This method will be illustrated later in a numerical example.

As Van Eeden (32) has remarked, the estimates obtained by this pooling technique are the maximum likelihood estimates of the P_i's under H_o. This fact can easily be seen by noting that if $p_i > p_{i+1}$, the maximum of the likelihood function under H_o occurs at the boundary of the feasible region, $P_i = P_{i+1} = P$, since

$$L = P_i^{Y_i} Q_i^{n_i - Y_i} \cdot P_{i+1}^{Y_{i+1}} Q_{i+1}^{n_{i+1} - Y_{i+1}}$$

$$= P^{Y_i + Y_{i+1}} Q^{n_i + n_{i+1} - (Y_i + Y_{i+1})}.$$

The expected number of responses at dose z_i, under H_o is thus estimated by $p_i'n$ and the usual chi-square goodness of fit criterion applied in this case leads to the following test statistic

$$\lambda = \sum_{i+1}^{k} \frac{[Y_i - p_i'n]^2}{p_i'n} + \frac{[(n - Y_i) - (1 - p_i')n]^2}{(1 - p_i')n}. \qquad (4)$$

If r ($\leq$k) groups were obtained by the pooling process described, then there are r different p_i''s. It can be seen that contributions to the calculated value of λ are made only by the pooled groups. The calculated λ is based on k - r degrees of freedom and, if no account were taken of the pooling procedure, this statistic would, be asymptotically

distributed as χ^2_{k-r}, under the special case of the null hypothesis, H_o': $P_1 = P_2 = \ldots = P_k$. This can be seen by noting the results of Bartholomew (3, p. 41) where it is shown that if the Y_i's are normally distributed with a common variance, σ^2, then estimates obtained by the pooling procedure described may be treated as though they were independent.

In order to derive the approximate distribution of $\lambda|H_o'$, allowance must be made for the probability of obtaining r different groups out of k initial groups by the described amalgamation process. Let P(r, k) denote this probability. The approximate probability that a calculated value of λ will exceed a given λ_o may therefore be written as

$$P_r(\lambda \geq \lambda_o) = \sum_{r=1}^{k-1} P(r, k) Pr(\chi^2_{k-r} \geq \lambda_o), \qquad (5)$$

or

$$Pr(\lambda \geq \lambda_o) = \sum_{r=1}^{k-1} P(r, k) \int_{\lambda_o}^{\infty} p(\chi^2_{k-r}) d\chi^2_{k-r}. \qquad (6)$$

In 1959, Miles (25) showed that $P(r, k) = |S_k^r|/k!$, if the groups are equally weighted, where S_k^r is the Stirling number of the first kind. A table of these numbers may be found in Miles's paper. The distribution in (6) is analogous to the $\bar{\chi}^2$ distribution defined by Bartholomew (3) in 1959. However, because a different hypothesis is of interest here, the degrees of freedom associated with the chi-square terms are not the same in (6) as in Bartholomew's $\bar{\chi}^2$. As a result

the distributions are not the same.

The upper 1 and 5 percentage points for λ for certain values of k are shown in Table 2. These points were calculated from (5) using a table of the Stirling numbers and the distribution function of χ^2 as tabulated by Pearson and Hartley (27). Interpolation was used to obtain the third significant figure of the percentage points and, as a result, there may be an error of one unit at that figure. The significance of a calculated λ as given in expression (4), may then be tested by the use of Table 2 for the specified values of k. This statistic is appropriate for testing H_o', which again can be seen to be conservative in the sense that for any other H_o'' under H_o,

$$\Pr[\lambda > \lambda_o | H_o'] \geq \Pr[\lambda > \lambda_o | H_o''].$$

Table 2. Approximate percentage points for λ

k	$\Pr(\lambda \geq \lambda_5) = .05$ λ_5	$\Pr(\lambda \geq \lambda_1) = .01$ λ_1
3	4.61	7.73
4	6.19	9.58
5	7.69	11.3
6	9.10	13.0
8	11.9	16.1
10	15.0	19.1

As a simple illustration of the calculations involved when the λ criterion is used to test for the monotonicity of $F(z)$, data from Carlson (9) used in the previous example have been altered. Suppose that (where $n = 30$, at each x) the following observations had been made.

Dose (coded)	Number dead
1	$Y_1 = 7$
2	$Y_2 = 5$
3	$Y_3 = 15$
4	$Y_4 = 21$
5	$Y_5 = 21$
6	$Y_6 = 18$

Since $Y_5 > Y_6$, these groups are averaged, and since $Y_4 > \frac{Y_5 + Y_6}{2} = 19.5$, Y_4, Y_5 and Y_6 are averaged. Y_3 is then less than $\frac{Y_4 + Y_5 + Y_6}{3} = 20$. Since $Y_1 > Y_2$, these two are also pooled. The p_i''s are therefore

$$p_1' = \frac{6}{30} = .20 \qquad p_4' = \frac{20}{30} = .667$$

$$p_2' = \frac{6}{30} = .20 \qquad p_5' = \frac{20}{30} = .667$$

$$p_3' = \frac{15}{30} = .50 \qquad p_6' = \frac{20}{30} = .667$$

where $p_1' \leq p_2' \leq \ldots \leq p_6'$. The calculated value of λ in this case is

$$\lambda = \frac{(7 - 6)^2}{6} + \frac{(24 - 23)^2}{24} + \frac{(6 - 5)^2}{6} + \frac{(25 - 24)^2}{24}$$

$$+ \frac{(21 - 20)^2}{20} + \frac{(10 - 9)^2}{10} + \frac{(21 - 20)^2}{20} + \frac{(10 - 9)^2}{10}$$

$$+ \frac{(22 - 20)^2}{20} + \frac{(20 - 18)^2}{10}$$

$$= 1.32$$

This value is less than the significance levels for k = 6 and the hypothesis of monotonicity of F(z) is accordingly accepted at $\alpha = .05$.

C. The Quantitative Response Case

1. A range test for monotonicity

Let the number of observations, n, be the same at each independent dose z_i (i = 1, . . ., k). Let the y_{ij}(j = 1, . . ., n) observations at each z_i be normally distributed with mean $E(y_{ij}) = Y_i$ and a common but unknown variance σ^2. The Y_i's are assumed to lie on the dose-response curve $Y_{ij} = F(z_i)$.

It is desired to test the hypothesis H_o: $Y_1 \leq Y_2 \leq \ldots \leq Y_k$ against the alternative hypothesis H_A: $Y_i > Y_{i_m}$ (i = 1, 2, . . ., k; m = 1, 2, . . ., k - i) for at least one i and m. As with the quantal response case, consider the special case H_o': $Y_1 = Y_2 = \ldots = Y_k$ under H_o. To test H_o against H_A, tables of the studentized range may be used. The criterion for a test of size α is to compare

$$\max (\bar{y}_i - \bar{y}_{i+m}) \text{ with } \frac{s}{\sqrt{n}} q_{2\alpha,\, k(n-1)}$$

where q is tabled by Pearson and Hartley (28), and s is the usual "within dose" estimate of σ, based on k (n-1) degrees of freedom. Since a one sided test is required in the present

situation, the 2α percentage points of q must be used; however, only the upper 5 and 1 percentage points for the studentized range are given in standard tables and the signifcance tests must, at present, be made at the 2.5 and 0.5 percentage points. Again, this test is conservative in the sense that the probability of rejecting H_o', and accepting H_A is greater under H_o' than for any other case under H_o.

2. A basis for an alternative test for monotonicity

Suppose that there are n_i observations at each independent dose z_i $(i = 1, \ldots, k)$. Let the y_{ij} $(j = 1, \ldots, n_i)$ observations at each z_i be independently and normally distributed with mean Y_i and a common unknown variance σ^2. It is assumed that the Y_i's lie on the curve $Y_i = F(z_i)$. Let

$$\sum_{i=1}^{k} n_i = n.$$

It is desired to test the hypothesis,

$$H_o: Y_1 \leq Y_2 \leq \ldots \leq Y_k$$

against the alternative

$$H_A: Y_i > Y_{i+m}, \text{ for at least one i and m}$$

$(i = 1, \ldots, k)$ $(m = 1, \ldots, k-i)$.

The likelihood function in this case is

$$L = (2\pi)^{-n/2}\sigma^{-n}e^{-n/2}.$$

The maximum likelihood estimator for σ^2 under the condition that $Y_1 \leq \ldots \leq Y_k$, as indicated by Bartholomew (3), is

is $s_o^2 = \frac{1}{n}\sum_{k=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - m_i)^2$ where the m_i are obtained by the following pooling process. That is, $m_i = \bar{y}_i$, if $\bar{y}_i \leq y_{i=1}$, whereas if $\bar{y}_i > \bar{y}_{i+1}$, then $m_i = m_{i+1} =$

$$\frac{\sum_{j=1}^{n_i} y_{ij} + \sum_{j=1}^{n_{i+1}} y_{(i+1)j}}{n_i + n_{i+1}}.$$

Starting with $\bar{y}_k$, this process is continued until $m_1 \leq m_2 \leq \ldots \leq m_r$, $r \leq k$. The maximum likelihood estimator for σ^2 under H_A can be seen to be the same as its unrestricted estimator,

$$s_A^2 = \frac{1}{n}\sum_{k=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_i)^2.$$

The likelihood ratio principle yields

$$\lambda = \frac{s_o^{-n}}{s_A^{-n}} = \frac{s_A^n}{s_o^n}.$$

Thus

$$\lambda^{2/n} = \frac{\sum_{k=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - m_i)^2}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_i)^2}$$

$$= \frac{\sum_{k=1}^{k}\sum_{j=1}^{n_i}[(y_{ij} - \bar{y}_i)^2 + (\bar{y}_i - m_i)^2]}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_i)^2}$$

$$= 1 + \frac{\sum_{i=1}^{k} n_i (\bar{y}_i - m)^2}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_i)^2} .$$

This leads to the consideration of

$$\lambda_k^* = \frac{\sum_{1}^{r} n_i(\bar{y}_i - m_i)^2}{(r-1)s_e^2} = \frac{s_b^2}{s_e^2} , \quad \text{where}$$

$$s_e^2 = \frac{\sum_{1}^{k}\sum_{1}^{n_i}(y_{ij} - \bar{y}_i)^2}{n-k} .$$

The distribution of λ_k^* under H_o can be shown to be closely related to Bartholomew's $\bar{F}_k$ criterion, the only difference being in the degrees of freedom of the F ratio. It has been shown (2) that

$$Pr(F_k \geq \gamma) = \sum_{2}^{k} P(r, k)\ Pr[F_{r-1,\ n-r} \geq \gamma],$$

and it can similarly be shown that

$$\Pr(\lambda' \geq \gamma) = \sum_{2}^{k} P(r, k) \Pr[F_{k-r,\, n-r} \geq \gamma]$$

where $P(r, k)$ is the probability of obtaining r almalgamated groups by the pooling procedure out of k groups. Percentage points for the distribution of λ' are, however, difficult to obtain and have not been tabulated. At the present time, the one tailed range test would be the recommended procedure for testing H_o against H_A in quantitative response assays.

D. Investigations of Similarity

1. Preliminary graphical investigation of similarity

The plotting of data on a graph as a first step in a statistical analysis is frequently of value in deciding on an appropriate method of analysis. Reasonable models and/or possible useful transformations may sometimes be seen more easily by this technique. A common problem in the usual analysis of bioassays is the determination of an appropriate response metameter (Y) which yields a linear relationship with the dose (z), or log dose (x). In assaying preparation S against preparation T by the usual methods, a Y must be found such that either

$$Y_S = \alpha + \beta_S z_S$$

$$Y_T = \alpha + \beta_T z_T \quad \text{(slope ratio model)} \qquad (7)$$

or

$$Y_S = \alpha_S + \beta x_S$$

$$Y_T = \alpha_T + \beta x_T$$

(parallel line model) (8)

are reasonable models to be employed.

For the above, or any other models based on similarity to be appropriate, $F_S(z_S)$ and $F_T(z_T)$ must be of the same functional form, that is $F_T(z_T) = F_S(\rho z_T)$, for all z_S, z_T. It may be difficult to tell whether this condition holds or not simply by looking at the data or even a plot of the data. If the condition definitely does not apply, then there is no need to search for a response metameter that will allow use of (7) or (8). A quick approximate check on the existence of similarity may be made by the following graphical procedure.

The observed responses are plotted against dose (z) for both preparations S and T. Using a French curve, a smooth monotonic curve is drawn by eye through the points for S and, independently for T, as shown in Figure 1. If these curves represent the dose-response relationships fairly well, a graph of equally effective doses of z_S and z_T may be obtained. An estimate of the z_S which is equivalent (in its ability to produce the response) to a specified z_T, is the z_S which corresponds to the response Y (z_T). A curve approximating the relationship between equally effective doses may be obtained by plotting the z_S's corresponding to equally spaced

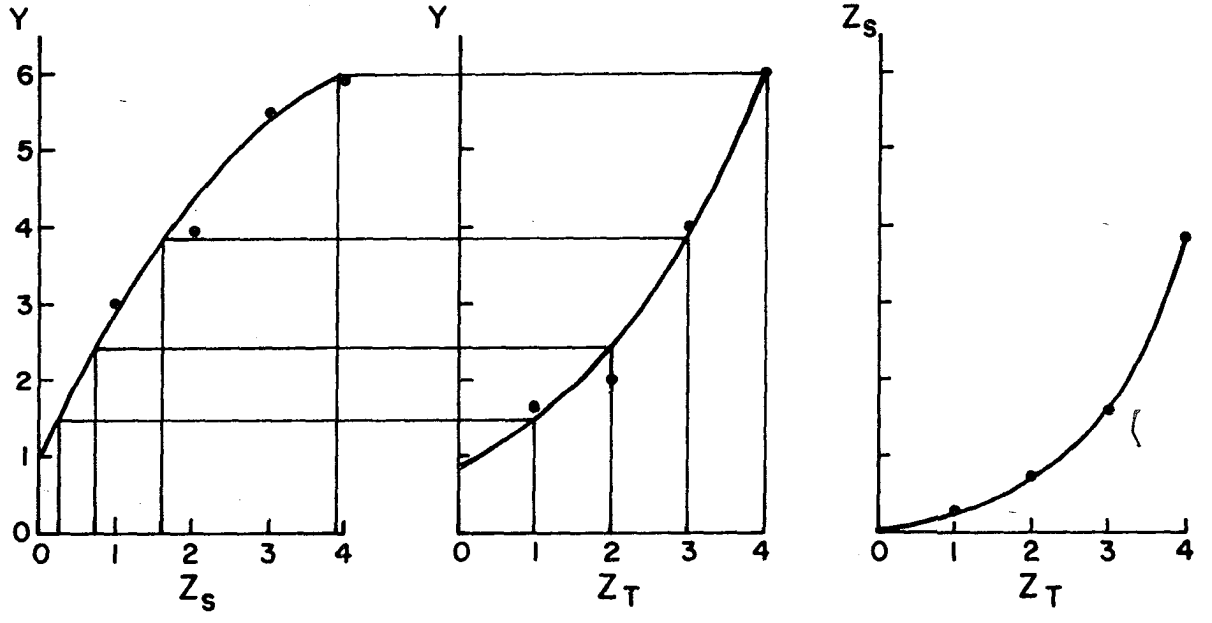

Figure 1. A graphical investigation of similarity

z_T's, as shown in the right hand graph in Figure 1. A linear appearance of this curve would tend to support the hypothesis of similarity, the slope of the line $z_S = \rho z_T$ being a rough estimate of relative potency. If however, this relationship is decidedly non-linear then the indication is that similarity does not apply, in which case no transformations of the data will allow use of a slope ratio or a parallel line assay model. In this case, although not of the simple $z_S = \rho z_T$ form, it might be desirable to find an appropriate model for the relationship and estimate its parameters.

2. Tests for similarity

If the condition of similarity applies, then, as discussed in Chapter I, relationships may always be expressed as $Y_S = F_S(x_S)$ and $Y_T = F_T(x_S - M)$ where x = log dose z, and M is log ρ. In the usual parallel line assay analysis, an appropriate linearizing response metameter is employed and the test for similarity is then based on deviations from the parallel line model. The following general test for similarity may be made (if monotonicity of the two curves is assumed) without assuming a specific form for F(x), by using the range tests given in Sections B and C.

Let there be n subjects exposed at each of the k_s doses of the standard preparation S and the k_T doses of the test preparation T. Under the hypothesis of similarity, the two observed log dose-response curves approximate to the same

monotonic function apart from their being horizontally displaced by an unknown amount M. Suppose that the observations on both preparations are amalgamated, by a trial horizontal translation of the observations on one of the preparations, to approximate a single monotonic function. Suppose, also, that after such a translation, the response observed at each dose is subtracted from the responses observed at every dose on its left. The maximum value obtained out of all differences taken in this manner may be noted. The particular translation(s) which yields the minimum value of this maximum difference is(are) the optimal translation(s) in this special snese. If, however, this minimum value exceeds the appropriate critical value of the range test for monotonicity, for $k_S + k_T$ samples of size n, then it seems reasonable to reject the hypothesis that the two sets of observations arose from monotonic functions of exactly the same shape, since combination of all observations, even in this optimal manner, would not allow acceptance of the hypothesis of monotonicity.

The following numerical example of this type of test is based on hypothetical data. Suppose that in a quantal response assay with equally spaced log doses (on the same scale) where 50 subjects were exposed at each dose, the number having responded at each dose were

$$Y_{S1} = 5 \quad Y_{T1} = 4$$
$$Y_{S2} = 38 \quad Y_{T2} = 20$$
$$Y_{S3} = 45 \quad Y_{T3} = 25$$

The minimum value of the maximum difference (taken in the manner described), that can be obtained by a horizontal translation of the observations, is 13, as illustrated in Figure 2. The appropriate value of $C_{6,\ .05}$ obtained from Table 1 is 1.83. Since 13 is greater than $1.83\sqrt{30} = 10.03$, the hypothesis of similarity would accordingly be rejected at $\alpha = .05$.

As would be expected, the hypothesis of similarity is rejected when a standard parallel line probit model is fitted to this data. Finney (15, p. 471-472) suggests procedures applicable in situations where tests of linearity are significant, but the assay is considered invalid, in any case, if similarity cannot be accepted. Other possible alternatives in such cases are discussed in Chapter IV.

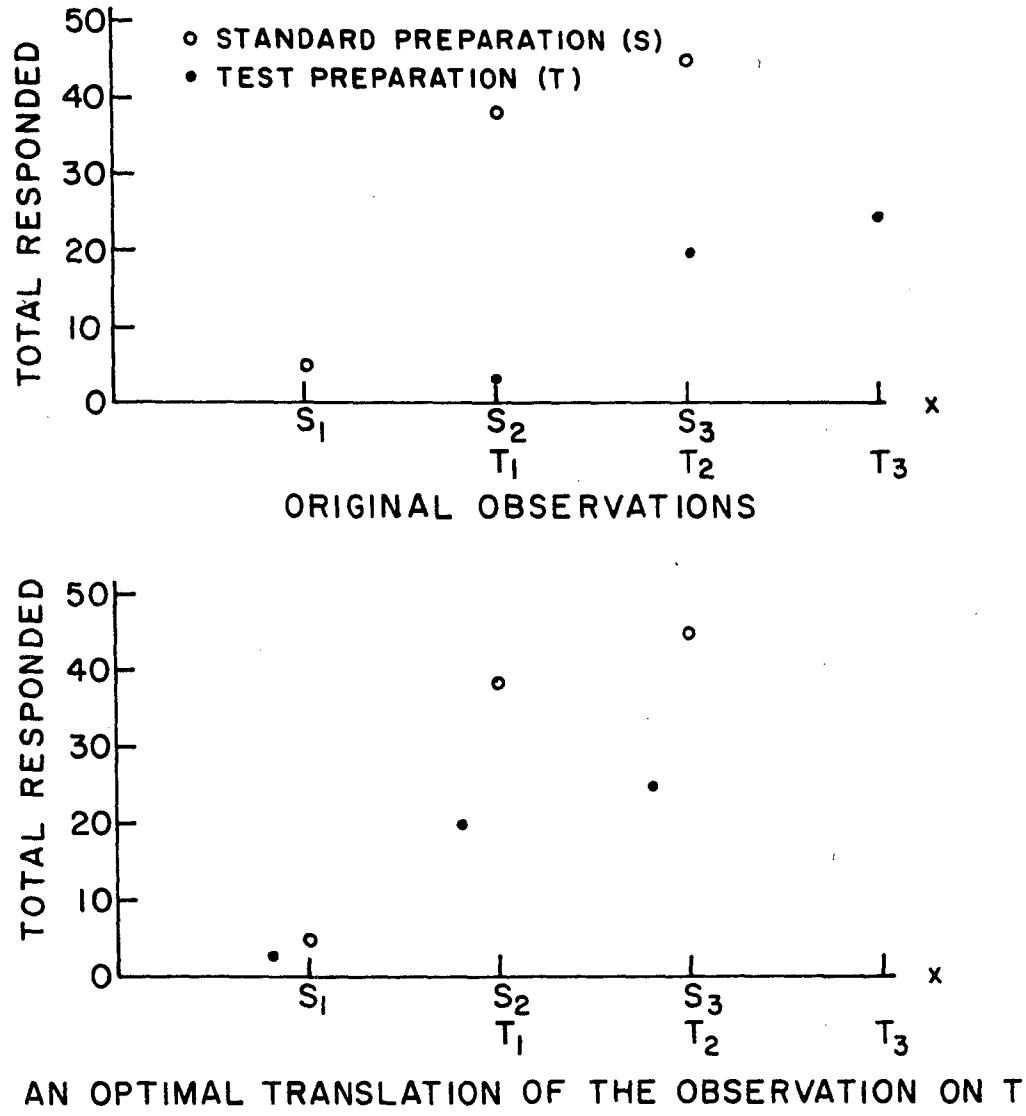

Figure 2. Translation of the observations on T

III. DISTRIBUTION-FREE ESTIMATION OF RELATIVE POTENCY IN QUANTAL RESPONSE BIOASSAYS

A. Some Distribution-Free Estimators of Relative Potency

Several methods have been developed for the estimation of the LD-50, that is, the median of the tolerance distribution, in quantal response assays. In some of these methods (e.g. probit, logit), the assumption is made that the tolerance distribution is of a known parametric form. Other methods, referred to by Finney (14, p. 364) as, " . . . the objective yet rapid methods . . .," do not require specification of an exact functional form for the tolerance distribution and are distribution-free in that respect. An estimate of the relative potency of two preparations may be determined by the ratio of two such LD-50 estimates (under the hypothesis of similarity). These distribution-free estimators are usually easy to compute and, although their properties are often difficult to evaluate, the methods have found application in many practical cases.

One of the simplest of the distribution-free techniques for estimating relative potency is that of linear interpolation. In this case, an estimate of the LD-50 is obtained for both the test and standard preparations by linearly interpolating between the two observed response percentages which bracket the fifty per cent points. This procedure has been recommended

by Doudoroff et al. (13) for use in toxicity studies on fish, for example. An obvious objection to this general method is, as Finney notes (14, p. 343), that it only uses part of the data (two doses from each preparation).

Other distribution-free techniques for providing estimates of the LD-50 and relative potency, such as the Dragstedt-Behrens, Reed-Muench, and Spearman-Karber methods, are discussed by Finney. In this review of the advantages and disadvantages of these and other methods, Finney (14, p. 364) states, with some reservations, that the Spearman-Karber procedure is perhaps the best of the objective yet rapid methods available for estimating relative potency and recommends that the Dragstedt-Behrens and Reed-Muench techniques be completely abandoned.

1. The Spearman-Karber estimator

Let n subjects be exposed at each of k doses of a preparation, and let y_i denote the number having responded at dose i. If the log dose levels (x_i) are x_1, $x_2 = x_1 + d$, . . ., $x_k = x_1 + (k-1)d$, the Spearman-Karber estimator of the LD-50 is

$$\bar{x} = p_1(x_1 - \tfrac{1}{2}d) + \sum_{i=1}^{k-1}(x_i + \tfrac{1}{2}d)(p_{i+1} - p_i) + (1-p_k)(x_k + \tfrac{1}{2}d)$$

$$= x_k + \tfrac{1}{2}d - d\sum_{i=1}^{k} p_i, \text{ where } p_i = \frac{y_i}{n}.$$

A symmetrical tolerance distribution is assumed, in which case the mean equals the median. Another necessary assumption is that the probability of a response is zero at all doses to the left of $x_1 - d$, and that the probability of a response is 1.00 at $x_k + d$ and higher doses. As Finney (14, p. 347) points out, the latter assumption may not be fulfilled in many situations. In such cases it is therefore worthwhile to consider alternative estimators.

B. An Alternative Distribution-Free Method for the Estimation of Relative Potency

As has been previously pointed out, under the condition of similarity, the log dose-response curves corresponding to the standard and test preparations have exactly the same shape, but are separated horizontally by an unknown distance M. In the quantal response case, for example, this implies that

$$F_S(x_S) = F_T(x_S - M),$$

where F_S and F_T are the distribution functions corresponding to the tolerance distributions for preparations S and T respectively. The problem of interest for bioassay purposes is that of estimating M, the horizontal displacement of the two monotonic curves, without assuming a parametric form for F.

It is rarely clear what criterion should be used for the determination of a distribution-free estimator. One possible criterion which might be considered in this case is the

"average" horizontal distance between the two observed log-dose response curves over the observed response range common to both curves. The problem of defining the average distance then arises. One such average may be obtained by an extension of the simple linear interpolation technique described in Section A, as follows:

Suppose that in a quantal response bioassay there are n_i subjects exposed at each log dose, $x_{Si}(i = 1, 2, \ldots, k_S)$, of the standard preparation S and each log dose, $x_{Tj}(j = 1, 2, \ldots, k_T)$, of the test preparation T. Let x_{Si} and x_{Tj} be equally spaced. Denote the number having responded at x_{Si} by y_i and similarly y_j at x_{Tj}, and let the corresponding proportions be $p_i = \frac{y_i}{n_i}$ and $p_j = \frac{y_j}{n_i}$.

An estimate of the value of x_T which corresponds to an observed p_{Si} may be found by linear interpolation between the two x_{Tj}'s which correspond to the two p_{Tj}'s (if there are such) which bracket the observed p_{Si}. The distance between an x_{Si} and the x_T estimate thus obtained provides an estimate of M. Estimates of M may be made in this manner for all values of p_{Si} which lie in the range of observed responses common to both preparations, and similarly for all values of p_{Tj} in the common response range. These separate estimates of M are not all independent, but even so the average of all of the estimates that can be obtained in this way provides an estimate of the average horizontal distance between the two response

curves.

1. Numerical example

Data for the following example is taken from an example used by Finney (14, p. 342) in which estimates of the relative potency computed by various methods were compared. The results of a quantal response assay were as shown below.

	Standard preparation			Test preparation		
coded log dose	number exposed	number killed	per cent	number exposed	number killed	per cent
-1	94	21	22.3	96	5	5.2
0	98	38	38.8	96	20	20.8
+1	96	64	66.7	97	56	57.7

These results are illustrated in Figure 3. The three horizontal lines in the figure are estimates, obtained in the manner described in this section, of the distance between the two response curves. The three calculated values are

$$1 + \frac{22.3 - 20.8}{57.7 - 20.8} = 1 + \frac{1.5}{36.9} = 1 + .041 = 1.041$$

$$\frac{38.8 - 20.8}{57.7 - 20.8} = \frac{18.0}{36.9} = 0.488$$

$$\frac{66.7 - 57.7}{66.7 - 38.8} = \frac{9.0}{27.9} = 0.323,$$

and the average of the these three estimates is

$$\hat{M} = 0.617.$$

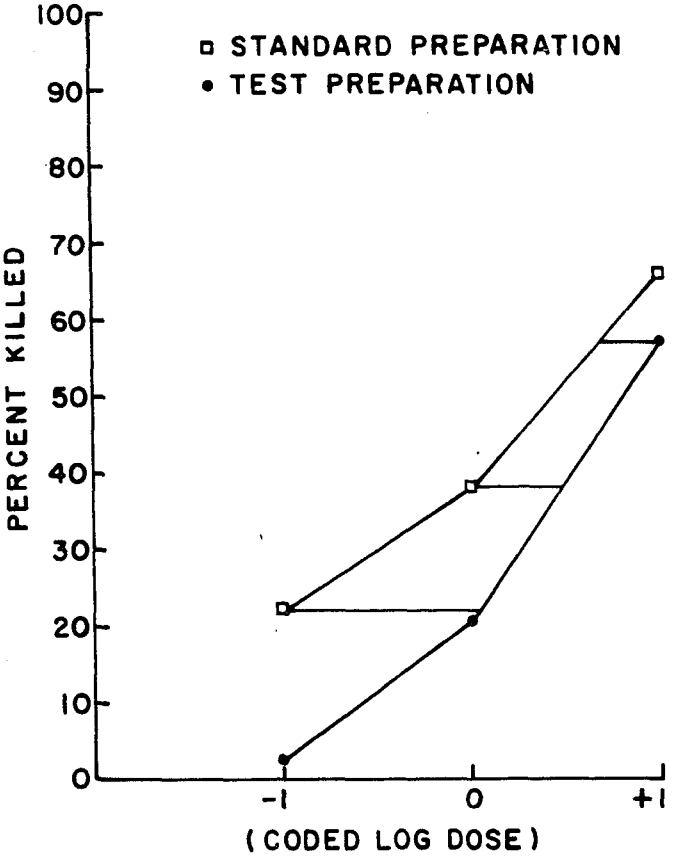

Figure 3. Results of a quantal response bioassay

2. Some empirical comparisons with other methods

It is difficult to compare distribution-free estimators other than on an empirical basis, as Finney has done, since in most cases their biases and variances are not known. In 1961, Brown (6) discussed these properties for the Spearman-Karber estimator in a hypothetical experiment in which an infinity of dose levels were available. A formula for the variance and bounds for the bias of this estimator are available only in the case when complete coverage of the response curves are obtained. This is not a common situation in bioassays. No formulas for the variance or bounds for the bias of the alternative "average distance" method discussed in this chapter are yet available. Some preliminary investigations are given in Appendix A.

As a partial evaluation of the average distance estimator it is of interest to compare estimates obtained by this method with those obtained by some other methods, both distribution-free and parametric, in some specific bioassays.

a. Example 1 For the data in the previous numerical example, Finney (14) has calculated estimates of M by several methods, five of which are shown below along with the average distance estimate.

Finney (14, p. 364) indicates that there is little doubt that the true value of M lies close to the results obtained by the probit and logit methods, and points out that all of the

Method	$\hat{M}$
Dragsted-Behrens	0.394
Reed-Muench	0.401
Spearman-Karber	0.441
probit	0.662
logit	0.634
average distance	0.617

distribution-free estimates were far too small. The Spearman-Karber estimate given was qualified in that it was obvious that one of its basic assumptions, previously discussed in Section A, was not valid. The estimate of M (0.617) calculated by the simple average distance method can be seen to be much closer to the probit and logit estimators, in this particular example, than any of the distribution-free estimates calculated by Finney.

b. Example 2. Burn et al. (8, p. 137)

Standard Preparation				Test Preparation			
log dose	number exposed	number killed	per cent	log dose	number exposed	number killed	per cent
.602	20	2	10	.778	20	7	35
.778	20	16	80	.954	20	19	95

The estimates of M (log ρ) in this case as obtained by three different methods are

average distance method	$\hat{M} = -0.122$
Spearman-Karber	$\hat{M} = -0.106$
probit	$\hat{M} = -0.104$

c. Example 3. Burn et al. (8, p. 145)

Standard Preparation				Test Preparation			
log dose	number exposed	number convulsed	per cent	log dose	number exposed	number convulsed	per cent
0	12	1	8	0	12	1	8
1	24	8	33	1	24	16	67
2	24	15	62	2	24	22	92
3	10	8	80	3	10	10	100

Estimates of M for the above assay as provided by three different estimators are

average distance method	$\hat{M} = 0.888$
Spearman-Karber	$\hat{M} = 0.840$
probit	$\hat{M} = 0.959$

In the last two examples, there is little difference between the Spearman-Karber and the average difference estimates.

The latter method does not require the assumption of symmetry nor the assumption that the entire response range be covered by the doses, which are necessary assumptions in the Spearman-Karber estimation technique. The closeness of

the approximation to a response curve, by a series of straight lines connecting the observed percentages, is probably the major factor in determining whether or not the average distance method will give a reasonable estimate in a particular bioassay. The choice between the methods will depend upon which of the above assumptions appears to be the more reasonable in any specific situation.

IV. SOME GENERALIZED BIOASSAY MODELS AND METHODS

A. The Condition of Similarity

As stated in Chapter I, the condition of similarity requires that the test preparation (T) must behave as though it were a dilution (or a concentration) of the standard preparation (S) in a diluent that is completely inert with respect to the response used. This implies that the ratio of equally effective doses $\rho = \frac{z_S}{z_T}$ is constant, within the range of experimentation. The relationship between the two response curves can therefore be written as

$$F_T(z_T) = F_S(\rho z_T).$$

The truth of the assumption of similarity (analogous to the assumption a lack of interaction between treatment effects in the analysis of variance) allows the relative potency, of S with respect to T, to be expressed as a single number within the range of observations. This condition is, therefore, of obvious convenience for the concise presentation and application of the results of a bioassay. Validity tests for similarity are, in fact, made as part of the statistical analysis when the standard slope ratio or parallel line assay models are employed. Cases frequently occur, however, when the assumption is violated and there is then no specified standard procedure for analysis available. Another difficulty which may arise when the number of observations is small is

that the usual validity tests may not be sufficiently powerful to detect any such invalidity except in extreme cases. Correspondingly erroneous inferences are, therefore, possible.

The need for a suitable procedure for cases where the assumption of similarity breaks down is further evidenced by the lack of uniformity in some of the instances which have been published, as shown below.

1. Ing, et al. (20, p. 90) proposed to provide an average relative potency estimate over a wide range of doses.
2. Gibbs, et al. (16, p. 408) computed an estimate of log ρ at the 50 per cent level with the qualification that, ". . . these ratios, although invalid, are nevertheless useful approximations . . ."
3. Grimshaw and D'Arcy (18, p. 262), in 1960, emphasized that there is no adequate method for "quantitative assessment" in this situation.

Two other published examples are by Bülbring and Wajda (7) and David and Fellowes (12). In all of the above studies linearity of a response metameter Y with log dose was noted, but the lines for different preparations were obviously not parallel.

It is the purpose of this chapter to show that methods for providing a "quantitative assessment" of two preparations, S and T, in just such situations can be obtained.

Attention has been restricted, in the first instance, to linear and quadratic dose-response models, since these appear likely to be the most common and important cases.

B. Relative Potency as a Function of Response

One method that could be used to present the results of a bioassay when there is a lack of similarity would be to express relative potency as a function of the response Y. In this manner an estimate of the ratio of equally effective doses as a specified response level can be obtained as will be shown for linear log dose-response relationships.

Suppose that $Y = \alpha_S = \beta_S x_S$ and $Y = \alpha_T + \beta_T x_T$ are appropriate for the log dose-response relationships, where x = log dose, $\bar{x}_S$ and $\bar{x}_T$ are coded to be zero, and the usual Gauss-Markov conditions hold. For a general response level Y_o, say, the logarithm of the ratio of equally effective doses may be expressed as $x_S - x_T$. Thus

$$M(Y) = \log \rho\ (Y) = x_S - x_T$$

$$= \frac{Y - \alpha_S}{\beta_S} - \frac{Y - \alpha_T}{\beta_T}$$

$$= \frac{\beta_T(Y - \alpha_S) - \beta_S(Y - \alpha_T)}{\beta_S\beta_T}$$

$$= \frac{\beta_T Y - \beta_T\alpha_S - \beta_S Y + \beta_S\alpha_T}{\beta_S\beta_T}$$

$$= \left(\frac{\beta_S \alpha_T - \beta_T \alpha_S}{\beta_S \beta_T} \right) + \left(\frac{\beta_T - \beta_S}{\beta_S \beta_T} \right) Y, \qquad (9)$$

which may be written as

$$A + BY. \qquad (10)$$

The corresponding relative potency would be

$$\rho(Y) = 10^{A+BY},$$

assuming 10 is the base to which the logarithms were taken.

Estimators for A and B can be obtained by substituting a_S, b_S, a_T, b_T into (9) where these are the usual best linear unbiased estimators of α_S, β_S, α_T and β_T, obtained by fitting straight lines to the two sets of data independently. Equation (10) could thus be used to estimate M for a given value of the response, Y_o.

It can be seen in expression (9) that if $\beta_S = \beta_T = \beta$, then

$$M(Y) = x_S - x_T = \frac{\alpha_T - \alpha_S}{\beta} = M,$$

a constant for all Y. It thus appears that the standard parallel line assay model is simply that special case of (9) for which, under the condition of similarity, $\beta_S = \beta_T$. In this case, a set of parallel lines is fitted to the data and the appropriate parameter estimates are substituted in (9) to estimate M.

If the errors about the linear regression models are assumed to be NID (o, σ^2), the approximate variance for $\hat{M}(Y_o)$ may be found by using the approximate formulas for the

variance of a ratio of two variables (when the coefficient of variation of the denominator is small.) Thus

$$\text{Var } \hat{M}(Y_o) = \text{Var } \left(\frac{Y_o - a_S}{b_S}\right) + \text{Var } \left(\frac{Y - a_T}{b_T}\right),$$

where

$$\text{Var } \left(\frac{Y_o - a_S}{b_S}\right) \doteq \left(\frac{Y_o - a_S}{b_S}\right)^2 \left[\frac{\sigma^2_{a_S}}{(Y_o - a_S)^2} + \frac{\sigma^2_{b_S}}{b_S^2}\right]$$

and

$$\text{Var } \left(\frac{Y_o - a_T}{b_T}\right) \doteq \left(\frac{Y_o - a_T}{b_T}\right)^2 \left[\frac{\sigma^2_{a_T}}{(Y_o - a_T)^2} + \frac{\sigma^2_{b_T}}{b_T^2}\right].$$

Therefore,

$$\text{Var } \hat{M}(Y_o) \doteq \frac{b_S^2\sigma^2_{a_S} + (Y_o - a_S)^2\sigma^2_{b_S}}{b_S^4} + \frac{b_T^2\sigma^2_{a_T} + (Y_o - a_T)^2\sigma^2_{b_T}}{b_T^4}$$

$$\doteq \frac{b_T^4[b_S^2\sigma^2_{a_S} + (Y_o - a_S)^2\sigma^2_{b_S}] + b_S^4[b_T^2\sigma^2_{a_T} + (Y_o - a_T)^2\sigma^2_{b_T}]}{b_S^4 b_T^4}$$

$$\doteq \frac{b_T^4\left[\frac{b_S^2}{n_S} + \frac{(Y_o - a_S)^2}{\Sigma x_S^2}\right]\sigma_S^2}{b_S^4 b_T^4}$$

$$+ \frac{b_S^4\left[\frac{b_T^2}{n_T} + \frac{(Y_o - a_T)^2}{\Sigma x_T^2}\right]\sigma_T^2}{b_S^4 b_T^4},$$

if $\sigma_S^2 \neq \sigma_T^2$

$$\pm\left\{\frac{b_T^4\left[\frac{b_S^2}{n_S}+\frac{(Y_o-a_S)^2}{\Sigma x_S^2}\right]+b_S^4\left[\frac{b_T^2}{n_T}+\frac{(Y_o-a_T)^2}{\Sigma x_T^2}\right]}{b_S^4 b_T^4}\right\}\sigma^2,$$

if $\sigma_S^2 = \sigma_T^2 = \sigma^2$.

A similar treatment of the slope ratio assay case can be made. Suppose that $Y = \alpha_S + \beta_S z_S$ and $Y = \alpha_T + \beta_T z_T$ are valid models for the dose-response relationships where z is the dose and the Gauss-Markov conditions apply. The ratio of equally effective doses as a function of Y may be expressed as

$$\rho(Y) = \frac{z_S}{z_T} = \left(\frac{Y-\alpha_S}{\beta_S}\right)\Big/\left(\frac{Y-\alpha_T}{\beta_T}\right)$$

$$= \left(\frac{\beta_T}{\beta_S}\right)\left(\frac{Y-\alpha_S}{Y-\alpha_T}\right). \qquad (11)$$

It can be seen that if $\alpha_S = \alpha_T$, then expression (11) becomes

$$\rho(Y) = \frac{z_S}{z_T} = \frac{\beta_T}{\beta_S} = \rho,$$

a constant for all Y. The ordinary slope ratio assay model is, therefore, the special case of (11) for which $\alpha_S = \alpha_T$. The requirement that $\alpha_S = \alpha_T$ is thus seen to be equivalent to the condition of similarity in this situation.

Estimates of the parameters may be obtained from the independently fitted regression lines, and a value of $\rho(Y_o)$ may be estimated for a specified Y_o by the Equation (11). Under the same error assumptions as before, an approximate formula for the variance of $\hat{\rho}(Y_o)$ may be obtained. Thus

$$\text{Var}\left(\frac{Y_o-a_S}{b_S}\right) \doteq \left(\frac{Y_o-a_S}{b_S}\right)^2\left[\frac{\sigma^2_{a_S}}{(Y_o-a_S)^2}+\frac{\sigma^2_{b_S}}{b_S^2}-\frac{2\ \text{cov}\ (a_S,b_S)}{b_S(Y_o-a_S)}\right]$$

and

$$\text{Var}\left(\frac{Y_o-a_T}{b_T}\right) \doteq \left(\frac{Y_o-a_T}{b_T}\right)^2\left[\frac{\sigma^2_{a_T}}{(Y_o-a_T)^2}+\frac{\sigma^2_{b_T}}{b_T^2}-\frac{2\ \text{cov}\ (a_T,b_T)}{b_T(Y_o-a_T)}\right].$$

Therefore,

$$\text{Var}\,\hat{\rho}(Y_o) \doteq \frac{\left(\frac{Y_o-a_S}{b_S}\right)^2}{\left(\frac{Y_o-a_T}{b_T}\right)^2}\left[\frac{\left(\frac{Y_o-a_S}{b_S}\right)^2\left[\frac{b_S^2\sigma^2_{a_S}+(Y_o-a_S)^2\sigma^2_{b_S}}{b_S^2(Y_o-a_S)^2}-\frac{2b_S(Y_o-a_S)\text{cov}(a_S,b_S)}{b_S^2(Y_o-a_S)^2}\right]}{\left(\frac{Y_o-a_S}{b_S}\right)^2}+\frac{\left(\frac{Y_o-a_T}{b_T}\right)^2\left[\frac{b_T^2\sigma^2_{a_T}+(Y_o-a_T)^2\sigma^2_{b_S}-2b_T(Y_o-a_T)\text{cov}(a_T,b_T)}{b_T^2(Y_o-a_T)^2}\right]}{\left(\frac{Y_o-a_T}{b_T}\right)^2}\right]$$

$$\doteq \left(\frac{b_T}{b_S}\right)^2\left(\frac{Y_o-a_S}{Y_o-a_T}\right)^2\left[\frac{b_S^2\sigma_{a_S}^2+(Y_o-a_S)^2\sigma_{b_S}^2-2b_S(Y_o-a_S)\mathrm{cov}(a_S,b_S)}{b_S^2(Y_o-a_S)^2}\right.$$

$$\left. + \frac{b_T^2\,\sigma_{a_T}^2+(Y_o-a_T)^2\sigma_{b_T}^2-2b_T(Y_o-a_T)\mathrm{cov}(a_T,b_T)}{b_T^2(Y_o-a_T)^2}\right]$$

$$\doteq \left(\frac{b_T}{b_S}\right)^2\left(\frac{Y_o-a_S}{Y_o-a_T}\right)^2\left[\frac{\left[\frac{b_S^2}{n_S}+\frac{(Y_o-a_S)^2}{\Sigma(z_S-\bar{z}_S)^2}-\frac{2b_S(Y_o-a_S)\bar{z}_S}{\Sigma(z_S-\bar{z}_S)^2}\right]\sigma_S^2}{b_S^2(Y_o-a_S)^2}\right.$$

$$\left. + \frac{\left[\frac{b_T^2}{n_T}+\frac{(Y_o-a_T)^2}{\Sigma(z_T-\bar{z}_T)^2}-\frac{2b_T(Y_o-a_T)\bar{z}_T}{\Sigma(z_T-\bar{z}_T)^2}\right]\sigma_T^2}{b_T^2(Y_o-a_T)^2}\right].$$

An expression of the relationship between ρ and Y may not be the best method of presentation of the results in situations where similarity does not hold. Another possible procedure is discussed in the next section.

C. Equivalent Log Dose Relationships

1. $F_S(x_S) = \alpha_S + \beta_S x_S$, $F_T(x_T) = \alpha_T + \beta_T x_T$

An alternative method, which might prove more convenient in some circumstances than expressing ρ as a function of response, would be to obtain an estimate of the functional relationship between equivalent (in their ability to evoke the response) doses of preparations S and T. Suppose that $Y = \alpha_S + \beta_S x_S$ and $Y = \alpha_T + \beta_T x_T$, where x = log dose z. Upon

equating these two functions, it follows that

$$\alpha_S + \beta_S x_S = \alpha_T + \beta_T x_T$$

and

$$x_S = \frac{\alpha_T - \alpha_S + \beta_T x_T}{\beta_S} \tag{12}$$

$$= \left(\frac{\alpha_T - \alpha_S}{\beta_S}\right) + \left(\frac{\beta_T}{\beta_S}\right) x_T, \tag{13}$$

which may be written as

$$x_S = C + Dx_T. \tag{14}$$

It can be noted that in terms of the z's, $z_S = Kz_T^D$, in this case, where $C = \log K$.

Equation (14), therefore, expresses the linear relationship between equally effective log doses of S and T. The usual linear regression estimators a_T, b_T, a_S, b_S for parameters α_T, β_T, α_S and β_S, respectively, may be obtained by independently fitting the two linear models, where the response errors about both lines are assumed to be NID(o, σ^2). These estimates may then be used to estimate parameters C and D in (14).

a. Reduction to usual model If $\beta_T = \beta_S = \beta$, then expression (13) reduces to $x_S = \frac{\alpha_T - \alpha_S}{\beta} + x_T$, and

$$x_S - x_T = \frac{\alpha_T - \alpha_S}{\beta} = C. \tag{15}$$

As before, the usual parallel line model is seen to be a special case of (13) when the condition of similarity is fulfilled.

b. Fiducial limits For a given value of x_T, say x_{oT}, the equivalent (in its ability to produce the response) x_S, say x_{oS}, may be estimated by

$$\hat{x}_{oS} = \frac{a_T - a_S - b_T x_{oT}}{b_S},$$

as shown in Equation (12). The fiducial limits for x_{oS} may be obtained by a suitable adaptation of Fieller's theorem (see e.g. Finney (15, p. 27)).

Suppose that, in a quantitative response assay, there are n_S observations available for preparation S, and n_T observations available for preparation T. Suppose also that $\sigma_S^2 = \sigma_T^2 = \sigma^2$. Let the variable x_{oS}, for a given x_{oT}, be defined as the solution of (12). Let a variable u be denoted by

$$u = a_S + b_S x_{oS} - a_T - b_T x_{oT}.$$

The expected value of u then is E(u) where

$$E(u) = \alpha_S + \beta_S x_{oS} - \alpha_T - \beta_T x_{oT} = 0.$$

The variable u is normally distributed with variance V(u) given by

$$V(u) = V(a_S) + x_{oS}^2 V(b_S) + V(a_T) + x_{oT}^2 V(b_T) + 2x_{oS}\mathrm{Cov}(a_S, b_S) + 2x_{oT}\mathrm{Cov}(a_T, b_T)$$

$$= \left[\frac{1}{n_S} + \frac{1}{n_T} + \frac{(x_{oS} - \bar{x}_S)^2}{\Sigma(x_S - \bar{x}_S)} + \frac{(x_{oT} - \bar{x}_T)^2}{\Sigma(x_T - \bar{x}_T)^2} \right] \sigma^2.$$

Upon replacement of σ^2 by its pooled estimate s^2 it can be seen that

$$\frac{(a_S + b_S x_{oS} - a_T - b_T x_{oT})}{s\sqrt{\frac{1}{n_S} + \frac{1}{n_T} + \frac{(x_{oS} - \bar{x}_S)^2}{\Sigma(x_S - \bar{x}_S)^2} + \frac{(x_{oT} - \bar{x}_T)^2}{\Sigma(x_T - \bar{x}_T)^2}}} \sim t_n,$$

where N is the degrees of freedom upon which the calculation of s^2 was based. This expression may be used to determine the values of x_{oS} which correspond to the appropriate significant t values. Thus

$$\frac{(a_S + b_S x_{oS} - a_T - b_T x_{oT})^2}{\left[\frac{1}{n_S} + \frac{1}{n_T} + \frac{(x_{oS} - \bar{x}_S)^2}{\Sigma(x_S - \bar{x}_S)^2} + \frac{(x_{oT} - \bar{x}_T)^2}{\Sigma(x_T - \bar{x}_T)^2}\right] s^2} = t^2,$$

where t is $t_{\alpha/2}$ with N degrees of freedom. By expanding the above expression and collecting terms, a quadratic equation in x_{oS} is obtained:

$$\left[b_S^2 - \frac{t^2 s^2}{\Sigma(x_S - \bar{x})^2}\right] x_{oS}^2 + 2\left[\frac{t^2 s^2 \bar{x}_S}{\Sigma(x_S - \bar{x}_S)^2} + b_S(a_S - a_T - b_T x_{oT})\right] x_{oS}$$

$$+ (a_S - a_T + b_T x_{oT})^2 - t^2 s^2 \left[\frac{1}{n_S} + \frac{1}{n_T} + \frac{\bar{x}_S^2}{\Sigma(x_S - \bar{x}_S)^2}\right.$$

$$\left. + \frac{(x_{oT} - \bar{x}_T)^2}{\Sigma(x_T - \bar{x}_T)^2}\right] = 0.$$

The solution of a quadratic $Fx_{oS}^2 + Gx_{oS} + H = 0$ may be written in the form

$$x_{oS} = -\frac{G}{2F} \pm \sqrt{\frac{(G/2)^2 - FH}{F^2}} .$$

In this case,

$$-\frac{G}{2F} = \frac{\left[b_S^2 - \frac{t^2 s^2}{\Sigma(x_S-\bar{x}_S)^2}\right]\bar{x}_S + b_S(a_T + b_T x_{oT} - \bar{y}_S)}{\left[b_S^2 - \frac{t^2 s^2}{\Sigma(x_S-\bar{x}_S)^2}\right]}$$

$$= \bar{x}_S + \frac{b_S(a_T + b_T x_{oT} - \bar{y}_S)}{\left[b_S^2 - \frac{t^2 s^2}{\Sigma(x_S-\bar{x}_S)^2}\right]} ,$$

and

$$\left(\frac{G}{2}\right)^2 - FH = \left[\left(\frac{t^2 s^2}{\Sigma(x_S-\bar{x})^2} - b_S^2\right)\bar{x}_S + b_S(\bar{y}_S - a_T - b_T x_{oT})\right]^2$$

$$-\left[b_S^2 - \frac{t^2 s^2}{\Sigma(x_S-\bar{x}_T)^2}\right]\left[(-b_S\bar{x}_S + \bar{y}_S - a_T - b_T x_{oT})^2\right.$$

$$\left. - t^2 s^2 \left(\frac{1}{n_S} + \frac{1}{n_T} + \frac{\bar{x}_S^2}{\Sigma(x_S-\bar{x}_S)^2} + \frac{(x_{oT}-\bar{x}_T)^2}{\Sigma(x_T-\bar{x}_T)^2}\right)\right]$$

$$= \left[b_S^2 - \frac{t^2 s^2}{\Sigma(x_S-\bar{x}_S)^2}\right]\left[\frac{1}{n_S} + \frac{1}{n_T} + \frac{(x_{oT}-x_T)^2}{\Sigma(x_T-\bar{x})^2}\right] t^2 s^2$$

$$+ \frac{t^2 s^2 (y_S - a_T - b_T x_{oT})^2}{\Sigma(x_S-\bar{x}_S)^2} .$$

Thus,

$$\sqrt{\frac{(G/2)^2 - FH}{F^2}} = \frac{ts}{\left[b_S^2 - \frac{t^2 s^2}{\Sigma(x_S-\bar{x}_S)^2}\right]}\sqrt{\left[b_S^2 - \frac{t^2 s^2}{\Sigma(x_S-\bar{x}_S)^2}\right]\left[\frac{1}{n_S} + \frac{1}{n_T} + \frac{(x_{oT} - \bar{x}_T)^2}{\Sigma(x_T-\bar{x}_T)^2} + \frac{(\bar{y}_S-a_T-b_T x_{oT})^2}{\Sigma(x_S-\bar{x}_S)^2}\right]}.$$

The two $(1-\alpha)$ per cent fiducial limits are, therefore,

$$x_{oS} = x_S + \frac{b_S(a_T + b_T x_{oT} - \bar{y}_S)}{\left[b_S^2 - \frac{t^2 s^2}{\Sigma(x_S-\bar{x}_S)^2}\right]} \pm \frac{ts}{\left[b_S^2 - \frac{t^2 s^2}{\Sigma(x_S-\bar{x}_S)^2}\right]}$$

$$\sqrt{\left[b_S^2 - \frac{t^2 s^2}{\Sigma(x_S-\bar{x}_S)^2}\right]\left[\frac{1}{n_S} + \frac{1}{n_T} + \frac{(x_{oT}-\bar{x}_T)^2}{\Sigma(x_T-\bar{x}_T)^2}\right] + \frac{(\bar{y}_S-a_T-b_T x_{oT})^2}{\Sigma(x_S-\bar{x}_S)^2}} \qquad (16)$$

If g is defined as

$$g = \left[\frac{t\sqrt{V(b_S)}}{b_S}\right]^2 = \frac{t^2 s^2}{b_S^2\,\Sigma(x_S-\bar{x}_S)^2}$$

then the expression for the limits may be written as

$$x_{oS} = \bar{x}_S + \frac{1}{(1-g)}\left\{\frac{(a_T + b_T x_{oT} - \bar{y}_S)}{b_S} \pm \frac{ts}{b_S^2}\sqrt{b_S^2(1-g)\left[\frac{1}{n_S} + \frac{1}{n_T} + \frac{(x_{oT}-\bar{x}_T)^2}{\Sigma(x_T-\bar{x}_T)^2}\right] + \frac{(\bar{y}_S-a_T-b_T x_{oT})^2}{\Sigma(x_S-\bar{x}_S)^2}}\right\}, \qquad (17)$$

and if g is small then a shorter approximate expression for the limits may be obtained as

$$x_{oS} = \bar{x}_S + \frac{a_T+b_T x_{oT}-\bar{y}_S}{b_S} \pm \frac{ts}{b_S}\sqrt{\frac{1}{n_S} + \frac{1}{n_T} + \frac{(x_{oT}-x_T)^2}{\Sigma(x_T-\bar{x}_T)^2} + \frac{(\bar{y}_S-a_T-b_T x_{oT})^2}{b_S^2\,\Sigma(x_S-\bar{x}_S)^2}} \quad . \tag{18}$$

In passing, it may be noted that the above formulas may be modified, as Finney (14, p. 472) shows for the standard parallel line assay, to include the quantal response (probit) case. The corresponding limits in this case are

$$x_{oS} = \bar{x}_S + \frac{1}{(1-g)}\left\{\frac{a_T+b_T x_{oT}-\bar{y}_S}{b_s} \pm \frac{t}{b_S^2}\sqrt{b_S^2(1-g)\left[\frac{1}{\Sigma nwx_S} + \frac{1}{\Sigma nwx_T} + \frac{(x_{oT}-\bar{x}_T)^2}{\Sigma_{Txx}}\right] + \frac{(\bar{y}_S-a_T-b_T x_{oT})^2}{b_S^2\,\Sigma_{Sxx}}}\right\},$$

where w is the probit weighting coefficient (14, p. 448) and Σ_{xx} is the weighted sums of squares. Mean values are also weighted and, in this case,

$$g = \frac{t^2}{b_S^2\,\Sigma_{Sxx}} \quad ,$$

where t is a normal deviate (1.96 for 95 per cent limits).

c. Numerical example The following example is based on a quantitative response assay reported by Ing et al. (20), in which one hundred mice were exposed at each of three doses

for two preparations. The standard preparation was atropine sulphate and the test preparation was designated as P_{10}. The response y (pupil size) was measured in divisions (7.7 divisions = 1 mm.). Doses z were measured in micrograms and the dose metameter x was $x = \log_2 z$, giving $x_S = 1, 2, 3$ and $x_T = 1, 2, 3$.

The original response data were not published in the paper and for the purposes of illustration here, the following modifications were made.

1. Approximate means of the observed responses at each dose were obtained from a graph.
2. An estimate of σ^2 was calculated from data given for a smaller, similar assay of atropine sulphate, where homoscedasticity was assumed (also assumed that $\sigma_S^2 = \sigma_T^2 = \sigma^2$).

The following response means were obtained

x_S	$\bar{y}_S$	x_T	$\bar{y}_T$
1	4.64	1	7.77
2	9.11	2	10.80
3	13.84	3	13.66.

The estimate of σ^2 was calculated to be $s^2 = 60.27$, which would have been based on 294 degrees of freedom in the original analysis.

The following quantities were computed to estimate the linear regressions for the two preparations separately.

$a_S = -.003$ $\qquad$ $a_T = 4.853$

$b_S = 4.600$ $\qquad$ $b_T = 2.945$

$\bar{y}_S = 9.197$ $\qquad$ $\bar{y}_T = 10.743$

$$\sum_1^3 (x_T - \bar{x}_T)^2 = \sum_1^3 (x_S - \bar{x}_T)^2 = 200,$$

where $n_S = n_T = 300$.

Since $s^2_{b_S} = s^2_{b_T} = s^2_b = \dfrac{s^2}{\Sigma(x-\bar{x})^2} = \dfrac{60.27}{200} = .3013 = (.549)^2$,

it can be seen that b_S and b_T are significantly different from zero and from each other. The estimates required in expression (13) are then

$$\hat{C} = \frac{a_T - a_S}{b_S} = \frac{4.854 + .003}{4.6} = 1.056$$

$$\hat{D} = \frac{b_T}{b_S} = \frac{2.945}{4.6} = 0.640$$

Therefore, the estimated relationship between equally effective log doses of the two preparations is

$$x_S = 1.056 + (.640)x_T. \qquad (19)$$

Transformation of this result into the original dose units yields

$$z_S = 2^{1.056} z_T^{.640} = 2.079\ z^{.640}.$$

The log dose (x_{oS}) of atropine sulphate equivalent (in its ability to produce the response under consideration) to an arbitrary dose of preparation $P_{10}(x_{oT})$, may be estimated from Equation (19). If, for example, $x_{oT} = 1.5$, then

$$\hat{x}_{oS} = 1.056 + (.640(1.5) = 1.056 + .96 = 2.016$$

and the corresponding dose is

$$\hat{z}_{oS} = 2^{2.016} = 4.04 \text{ micrograms.}$$

In order to obtain 95 per cent fiducial limits, the following value was calculated:

$$g = \frac{t^2 s^2}{b_S^2 \Sigma(x_S - \bar{x}_S)^2} = \frac{(4)(60.27)}{(4.6)^2(200)} = .057.$$

Use of Equation (16) yields the following limits for x_{oS}

$$x_{oS} = 2 + \frac{.33810}{19.955} \pm \frac{15.52}{19.955}\sqrt{(19.955)(.006667 + \frac{0.25}{200})} + \frac{.0054}{200}$$

$$= 2.0169 \pm .3092$$

$$= 1.71, 2.33$$

Since g was found to be fairly small, the use of the simpler, more approximate limit formula (18) may be considered. In this case the calculated limits are

$$x_{oS} = 2 + \frac{.0735}{4.6} \pm \sqrt{\frac{15.52}{4.6} \quad .007917 + \frac{.0054}{4232}}$$

$$= 2.016 \pm .300$$

$$= 1.72, 2.32 ,$$

and by comparison with the limits from the exact formula (16), it appears that limits found by the simpler approximate

formula are satisfactory in this case.

d. Preliminary design considerations Finney (14, p. 175) has given a list of requirements of a good parallel line assay design with the aim of reducing the distance between the fiducial limits. It is accordingly of interest here to list corresponding desiderata in the general non-parallel line situation. Consideration of Equation (17) shows that, in general, it would be desirable to choose a design such that

1. $\Sigma(x_T - \bar{x}_T)^2$ and $\Sigma(x_S - \bar{x}_S)^2$ are both large
2. $(\frac{1}{n_S} + \frac{1}{n_T})$ is small
3. $\frac{st}{b_S}$ is small

These requirements are the same as in the usual case discussed in detail by Finney, with the exception that b_S (instead of a pooled b) appears in the denominator of C.

It can also be seen from (17) that it is desirable that $|a_T + b_T x_{oT} - \bar{y}_S|$ and $|x_{oT} - \bar{x}_T|$ be small. The choice of doses may, therefore, also be influenced by the particular values of x_{oT} which are of most interest to the experimenter. Design considerations in situations where the condition of similarity is not assumed constitute an area where further research could be done. Specific design requirements will vary with many factors including purpose of the assay and the particular field of application.

2. Other models

Situations may arise in the analysis of bioassays where a linear model appears to be adequate for only one of the preparations, in which case the employment of a quadratic model for the other preparation may be a useful alternative. Two cases (a) and (b) below, may be distinguished according to whether the linear model is applicable to the standard preparation or to the test preparation.

a. $F_S(x_S) = \alpha_S + \beta_S x_S$, $F_T(x_T) = \alpha_T + \beta_T x_T + \gamma_T x_T^2$

Let $Y = \alpha_S + \beta_S x_S$, as in section 1. Suppose that a quadratic model is appropriate for the log dose-response curve for T, that is $Y = \alpha_T + \beta_T x_T + \gamma_T x_T^2$, where the error assumptions are as before. In this case it can be seen that

$$x_S = \frac{Y-\alpha_S}{\beta_S} = \frac{\alpha_T+\beta_T x_T+\gamma_T x_T^2-\alpha_S}{\beta_S}$$

$$=\left(\frac{\alpha_T-\alpha_S}{\beta_S}\right) + \left(\frac{\beta_T}{\beta_S}\right)x_T + \left(\frac{\gamma_T}{\beta_S}\right) x_T^2$$

$$= J + Kx_T + Lx_T^2$$

The above expression may be used to estimate the x_{oS} equivalent to a specified x_{oT}, where the usual regression estimators, a_S, b_S, a_T, b_T, c_T, are substituted for α_S, β_S, α_T, β_T, γ_T.

Fiducial limits for x_{oS} may be obtained in a manner similar to that previously described. The expression for the limits in this case is

$$\bar{x}_{oS} = \bar{x}_S + \frac{b_S(a_T+b_T x_{oT}+c_T x_{oT}^2-\bar{y}_S)}{\left[b_S^2 - \frac{t^2 s^2}{\Sigma(x_S-\bar{x}_S)^2}\right]} \pm \frac{ts}{\left[b_S^2 - \frac{t^2 s^2}{\Sigma(x_S-\bar{x}_S)^2}\right]}$$

$$\sqrt{\left[b_S^2 - \frac{t^2 s^2}{\Sigma(x_S-\bar{x}_S)^2}\right][V] + \frac{(\bar{y}_S-a_T-b_T x_{oT}-c_T x_{oT}^2)}{\Sigma(x_S-\bar{x}_S)^2}},$$

where $V = \frac{1}{s^2}\left[\hat{Var}(a_S) + \hat{Var}(a_T) + x_{oT}^2\hat{Var}(b_T) + x_{oT}^4\hat{Var}(c_T)\right.$

$$\left. + 2x_{oT}\hat{Cov}(a_T,b_T) + 2x_{oT}^2\hat{Cov}(a_T,c_T) + 2x_{oT}^3\hat{Cov}(b_T,c_T)\right]$$

and s^2 is the pooled estimate of σ^2, as before. Estimates for the variances and covariances in V may be obtained from the two variance-covariance matrices used in estimating the regression parameters.

b. $\underline{F_S(x_S) = \alpha_S + \beta_S x_S + \gamma_S x_S^2,\ F_T(x_T) = \alpha_T + \beta_T x_T}$ If $Y = \alpha_S + \beta_S x_S + \gamma_S x_S^2$ and $Y_T = \alpha_T + \beta_T x_T$, it can be seen that

$$\alpha_S + \beta_S x^2 + \gamma_S x_S^2 = \alpha_T + \beta_T x_T$$

and

$$x_S = \frac{-\beta_S \pm \sqrt{\beta_S^2 - 4\gamma_S(\alpha_S-\alpha_T-\beta_T x_T)}}{2\gamma_S}, \qquad (20)$$

which is an expression of the relationship between equally effective log doses of S and T in terms of x_S. The above expression may be used to provide an estimate of the x_{oS} equivalent to a specified x_{oT}, when, as before, the appropriate parameter estimates are substituted in (20).

Fiducial limits for x_{oS} are more difficult to compute, in this case, in that solving a quartic equation is involved.

As before, let a variable u be defined by

$$u = a_S + b_S x_{oS} + c_S x_{oS}^2 - a_T - b_T x_{oT}.$$

This variable is normally distributed with an expected value of zero and variance

$$V(u) = V(a_S) + x_{oS}^2 V(b_S) + x_{oS}^4 V(c_S) + 2x_{oS}\mathrm{Cov}(a_S,b_S)$$

$$+ 2x_{oS}^2\mathrm{Cov}(a_S,b_S) + 2x_{oS}^3\mathrm{Cov}(b_S,c_S) + V(a_T) + x_{oT}^2 V(b_T)$$

$$+ 2x_{oT}\mathrm{Cov}(a_T,b_T)$$

$$= \sigma^2 (v_{11_S} + x_{oS}^2 v_{22_S} + x_{oS}^4 v_{33_S} + 2x_{oS} v_{12_S}$$

$$+ 2x_{oS}^2 v_{13_S} + 2x_{oS}^3 v_{23_S} + v_{11_T} + x_{oT}^2 v_{22_T} + 2x_{oT} v_{12_T}),$$

if $\sigma_S^2 = \sigma_T^2 = \sigma^2$. The v's can be seen to be the usual elements of the inverse of the variance-covariance matrices used in determining the estimates of the parameters of the individual regression models.

The values of x_{oS} which correspond to the critical t values may be found in the same manner as before. Thus,

$$\frac{(a_S+b_S x_{oS}+c_S x_{oS}^2-a_T-b_T x_{oT}^2)}{s^2(v_{11_S}+x_{oS}^2 v_{22_S}+x_{oS}^4 v_{33_S}+2x_{oS} v_{12_S}+2x_{oS}^2 v_{13_S}+2x_{oS}^3 v_{23_S}+v_{11_T}}$$

$$\overline{+x_{oT}^2 v_{22_T}+2x_{oT} v_{12_T})} = t^2,$$

where t is $t_{\alpha/2}$ with the degrees of freedom corresponding to the degrees of freedom used in calculating s^2. Solving for x_{oS} in the above expression results in the following quartic

equation in x_{oS}

$$[c_S^2 - t^2s^2v_{33_S}]x_{oS}^4 + [2b_Sc_S - 2t^2s^2v_{23_S}]x_{oS}^3$$

$$+ [b_S^2 + 2a_Sc_S - 2a_Tc_S - 2b_Tc_Sx_{oT} + t^2s^2v_{22_S} + 2t^2s^2v_{13_S}]x_{oS}^2$$

$$+ [2a_Sb_S - 2a_Tb_S - 2b_Sb_Tx_{oT} - 2t^2s^2v_{12_S}]x_{oS}$$

$$+ [a_S^2 + a_T^2 + b_T^2x_{oT}^2 - 2a_Sa_T - 2a_Sb_Tx_{oT} + 2a_Tb_Tx_{oT}$$

$$+ t^2s^2(v_{11_S} + v_{11_T} + 2x_{oT}v_{12_T} + x_{oT}^2v_{22_T})]$$

$$= 0.$$

Methods for finding the roots of a quartic equation are available in most standard textbooks on the theory of equations. Adams (1) has designed a nomogram which may be of help in finding a quick approximate solution. Williams (33, p. 108-109) has discussed the interpretation of possible solutions of a quartic equation which arose in a similar context, that of determining fiducial limits in an inverse estimation problem.

If a quadratic model is employed for both preparations it is again necessary to solve a quartic equation to obtain the fiducial limits. Apart from the purely algebraic computations involved no differences in the general procedure arise in this case.

It is interesting to note that Bliss (5), in 1957, suggested the fitting of parallel quadratic functions, as an

alternative procedure to the parallel straight line model, when curvature is apparent which cannot satisfactorily be removed by a transformation. The decision as to whether a linear or quadratic model should be used in such an analysis is a special case of a problem in sequential model building, considered in a paper by Larson and Bancroft (22), in 1963, in which rules for making such a decision are examined in detail.

Bliss (5) did not consider the possibility of violating the condition of monotonicity when a quadratic model is used. In such cases, if the model $\gamma = \alpha + \beta x_i + \gamma x_i^2$ ($i = 1, \ldots, k$) is fitted, then two linear restrictions which must be imposed on the parameters are

$$\beta + 2\gamma x_1 \geq 0$$

$$\beta + 2\gamma x_k \geq 0$$

These restrictions insure that the first derivative is non-negative and therefore that the function is non-decreasing between x_1 and x_k. The least squares fitting of polynomial function with linear constraints on the parameters is a problem in quadratic programming. Lewish (23) has considered this general problem in a study of linear estimation in convex parameter spaces. A numerical example of fitting a non-decreasing quadratic function by least squares, based on theoretical developments given by Lewish, appears in Appendix B. This method is recommended for use whenever a quadratic model is employed to represent a response function in a bioassay.

V. SUMMARY

The standard statistical methods for the analysis and interpretation of bioassays are based on a number of assumptions concerning the relevance of the mathematical model to the behavior of the experimental material. The usefulness of such methods depends upon the appropriateness of these assumptions in a particular bioassay. Situations arise, in bioassay studies, where some of the usual assumptions are not justified so that the standard statistical methods are inapplicable. This dissertation is concerned with the investigation of some of the basic assumptions used in bioassays, and the development of procedures applicable when they are not fulfilled.

One basic assumption, which must be considered in the statistical analysis of all bioassays, is that the dose-response function, $F(z)$, is monotonic. Tests for monotonicity, where the parametric form of $F(z)$ is not specified, are developed in Chapter II for both the quantal and quantitative response cases.

Another assumption which is frequently made is that the response function is or can be made essentially linear, with respect to dose or log dose, by an appropriate transformation of the response variable. Standard tests of this assumption are available in the literature. When no exact parametric form, such as a linear model, for $F(z)$ is assumed, an esti-

mate of relative potency may still be made. This estimation problem is examined in Chapter III.

For the relative potency of one preparation with respect to another to be a constant value, the condition of similarity must be valid. This condition requires that the test preparation must behave as though it were a dilution (or a concentration) of the standard preparation in an inert diluent. Tests for the hypothesis of similarity based on an assumed parametric form for $F(z)$ are available for the usual models employed in bioassays. General methods for investigating similarity, without assuming an exact form for $F(z)$, are given in Chapter II.

Although all of the standard statistical procedures available for the analysis of bioassays are based on the assumption of similarity, many cases arise in practice in which this convenient assumption is violated. Examples of this situation in several bioassay studies have been cited. By the usual statistical procedures an assay is considered "invalid" if the hypothesis of similarity is rejected. Efficient alternative procedures for analysis in such cases are not generally available. Methods for the analysis and presentation of bioassay results when the assumption of similarity is not necessarily appropriate have been developed in Chapter IV.

Depending upon the purpose of a particular assay, the "general" methods of analysis in Chapter IV may be of use even in cases where the hypothesis if similarity is not

rejected. At the expense of conciseness of presentation (a single value $\hat{\rho}$), an estimate of the equivalent dose relationship may provide a better representation of the true situation. The criteria for determining the optimum analysis and presentation of bioassay results, particularly with regard to the comparison of the standard and general methods of analysis, is an area which requires further investigation.

VI. LITERATURE CITED

1. Adams, D. P. The quintic hypernom for the equation $x^5 + Ax^3 + Bx^2 + Cx + D = 0$. Journal of Mathematics and Physics 22: 78-92. 1943.

2. Bartholomew, D. J. Ordered tests in the analysis of variance. Biometrika 48: 325-332. 1961.

3. _____. A test of homogeneity for ordered alternatives. Biometrika 46: 36-48. 1959.

4. _____. A test of homogeneity for ordered alternatives. II. Biometrika 46: 328-335. 1959.

5. Bliss, C. I. Bioassay from a parabola. Biometrics 13: 35-50. 1957.

6. Brown, W. B. Some properties of the Spearman estimator in bioassay. Biometrika 48: 293-302. 1961.

7. Bülbring, E. and Wajda, I. Biological comparison of local anaesthetics. Journal of Pharmacology and Experimental Therapeutics 85: 78-84. 1945.

8. Burn, J. H., Finney, D. J. and Goodwin, L. G. Biological standardization. Second edition. Oxford University Press. London. 1950.

9. Carlson, C. A. Bottom fauna of the Mississippi River near Keokuk, Iowa, with particular reference to possible control of Ephemeroptera and Trichoptera. Unpublished Ph.D. thesis. Ames, Iowa, Library, Iowa State University of Science and Technology. 1963.

10. Cochran, W. G. Stochastic approximation to the median effective dose in bioassay. Unpublished paper presented at a central regional meeting of the Institute of Mathematical Statistics, Madison, Wisconsin, June 14, 1963. Department of Statistics, Harvard University, Cambridge, Mass.

11. Cox, C. P. Statistical methods in bioassay. Unpublished notes (mimeographed). Statistical Laboratory, Iowa State University of Science and Technology. Ames, Iowa. 1962.

12. David, A. and Fellowes, K. P. Some pharmacological properties of 5-chloro-2, 4-disulphamltoluene "disamide" an orally active diuretic agent. Journal of Pharmacy and Pharmacology 12: 65-73. 1960.

13. Doudoroff, P., Anderson, B. G., Burdick, G. E., Galtsoff, P. S., Hart, W. B., Patrick, R., Strong, E. F., Surber, E. W. and Van Horn, W. M. Bio-assay methods for the evaluation of acute toxicity of industrial wastes to fish. Sewage and Industrial Wastes 23: 1380-1397. 1951.

14. Finney, D. J. The design and analysis of an immunological assay. Acta Microbiologica (Hungaricae) 6: 341-368. 1959.

15. _____. Statistical method in biological assay. Hafner Publishing Company. New York. 1952.

16. Gibbs, M. B., Laffer, N. C., Dunne, C. J. and Akeroyd, J. H. Quantitative hemagglutination inhibition studies of blood group substances. II. Characterization of anti-A isohemagglutinins by their behavior with blood group A substances. Journal of Immunology 87: 405-414. 1961.

17. Goldberg, L. J. A visual approach to bio-assay. In De Jonge, H., ed. Quantitative methods in pharmacology. pp. 172-210. North-Holland Publishing Company. Amsterdam. 1962.

18. Grimshaw, J. J. and D'Arcy, P. F. Some problems arising in drug standarisation and the selective screening of compounds of potential pharmacological interest. In De Jonge, H., ed. Quantitative methods in pharmacology. pp. 248-263. North-Holland Publishing Company. Amsterdam. 1962.

19. Hartley, H O. The application of mathematical programming to statistical least squares estimation with constraints. (Mimeographed). Iowa State University of Science and Technology Statistical Laboratory Technical Report No. 1. 1963.

20. Ing, H. R., Dawes, G. S. and Wajda, I. Synthetic substitutes for atropine. Journal of Pharmacology and Experimental Therapeutics 85: 85-102. 1945.

21. Jerne, N. K. and Wood, E. C. The validity and meaning of the results of biological assays. Biometrics 5: 273-299. 1949.

22. Larson, H. J. and Bancroft, T. A. Sequential model building for prediction in regression analysis. I. Annals of Mathematical Statistics 34: 462-479. 1963.

23. Lewish, W. T. Linear estimation in convex parameter spaces. Unpublished Ph.D. thesis. Library, Iowa State University of Science and Technology. Ames, Iowa. 1963.

24. Mantel, N. and Hilgar, A. G. Effects of nonlinearity on subjective and objective analyses of hormone assay data. U. S. Department of Health, Education, and Welfare Public Health Service Cancer Chemotherapy Reports No. 11:61-68. [1960].

25. Miles, R. E. The complete amalgamation into blocks, by weighted means, of a finite set of real numbers. Biometrika 46: 317-327. 1959.

26. Pearson, E. S. The probability integral of the range in samples of n observations from a normal population. I. Forward and tables. Biometrika 32: 301-308. 1942.

27. _____ and Hartley, H. O. Biometrika tables for statisticians. Volume 1. Cambridge University Press. Cambridge, England. 1954.

28. _____ and _____. Tables of the probability integral of the studentized range. Biometrika 33: 89-99. 1943.

29. Thompson, W. R. On the use of parallel or nonparallel systems of transformed curves in bioassay: illustration in the quantitative complement-fixation test. Biometrics 4: 197-210. 1948.

30. Tukey, J. W. Comparing individual means in the analysis of variance. Biometrics 5: 99-114. 1949.

31. Van Eeden, C. A class of tests for the hypothesis that k parameters $\theta_1, \ldots, \theta_k$ satisfy the inequalities $\theta_1 \leq \ldots \leq \theta_k$. Statistical Department of the Mathematical Centre, Amsterdam. Report SP 65. 1960.

32. Van Eeden, C. On distributionfree bio-assay. In De Jonge, H., ed. Quantitative methods in pharmacology. pp. 206-209. North-Holland Publishing Company. Amsterdam. 1962.

33. Williams, E. J. Regression analysis. John Wiley and Sons. New York. 1959.

VII. ACKNOWLEDGMENTS

The author expresses his sincere appreciation to Professor C. Philip Cox for his many helpful suggestions and constant guidance throughout the course of this study.

Thanks are also given to the National Institutes of Health for their financial support of the research.

VIII. APPENDIX A

In Chapter III a simple "average distance" estimator of log ρ (relative potency), in quantal response bioassays, was described for situations where the hypothesis of similarity is accepted and no parametric form for the cumulative log tolerance distribution, F(x), is specified. Information about the properties of this (or any other distribution-free estimator) is difficult to obtain. Some preliminary considerations of the variance and expectation of this estimator are presented in this appendix.

A. Variance

Suppose that the number of subjects exposed at each x_{Si} ($i = 1, 2, \ldots, k$) of the standard preparation is n_{Si}. Let the number of responses, y_{Si}, which occur at each independent x_{Si}, be binomially distributed with the true probability of a response equal to $F(x_{Si})$. Let, also, $p_{Si} = \frac{y_{Si}}{n_{Si}}$. Similarly define $p_{Tj} = \frac{y_{Tj}}{n_{Tj}}$ for each x_{Tj} ($j = 1, 2, \ldots, k$) of the test preparation. The two response functions, $F_S(x_{Si})$ and $F_T(x_{Tj})$, are assumed to be of the same, but unknown, shape, and are horizontally separated by an unknown amount, M.

An estimate, m, of the quantity, M, may be made at every x_{Si}, where p_{Si} is in the range of responses common to both preparations. For each such x_{Si}, the corresponding m is

determined by the distance between x_{Si} and x_T, where x_T is estimated by simple linear interpolation between the two p_{Tj}'s which bracket p_{Si}. Let the two corresponding x_T's be designated as x_{Tj} and $x_{T(j+1)}$. An estimate of m may similarly be obtained for every p_{Tj} in the common response range. This procedure is described and illustrated in Chapter III.

If h such estimates are obtained, the average distance estimator is

$$\hat{M} = \frac{1}{h} \sum_{r=1}^{h} m_r.$$

The variance of $\hat{M}$ is

$$\text{Var } (\hat{M}) = \frac{1}{h^2} \left[\sum_r \text{Var } (m_r) + 2 \sum_{r \neq s} \text{Cov } (m_r, m_s)\right], \text{ where}$$

a particular m_r may be written as

$$m_r = \left[(x_{Tj} - x_{Si}) + \frac{p_{Si} - p_{Tj}}{p_{T(j+1)} - p_{Tj}}\right].$$

The variance of m_r involves the variances and covariance of the two terms within the brackets, and variance of $(x_{Tj} - x_{Si})$ depends upon the shape of F(x) in the range of experimentation. Therefore, the variance of an individual m_r, and hence the variance of $\hat{M}$, cannot be determined without further information about F(x). With further investigation, however, it may be possible to establish bounds for the variance of $\hat{M}$ by making certain assumptions about F(x) relative to the spacing of the x's.

B. Expectation

Consider again, an individual estimate,

$$m_r = (x_{Tj} - x_{Si}) + \frac{p_{Si} - p_{Tj}}{p_{T(j+1)} - p_{Tj}}.$$

The expectation of the above fraction obviously is between 0 and 1. The expectation of $(x_{Tj} - x_{Si})$ is actually a conditional expectation,

$$E[(x_{Tj} - x_{Si})|A]$$

where A is the event that $p_{Tj} \leq p_{Si} \leq p_{T(j+1)}$.
It follows, therefore, that

$$E[(x_{Tj} - x_{Si})|A] = \Sigma[x_{Tj} - x_{Si}]\ Pr[(x_{Tj} - x_{Si})|A],$$

where the summation is taken over all possible values of $(x_{Tj} - x_{Si})$ for which $Pr[(x_{Tj} - x_{Si})|A]>0$.

The above probabilities depend upon the shape of F(x) in the area in which the x_{Si}'s and x_{Tj}'s were chosen. As is the case with the variance, it would be impossible to establish bounds for the bias of $\hat{M}$ without making some assumptions about the shape of F(x) in this range. Although a number of such assumptions could be made, it appears difficult to define any one which would not be very nearly equivalent to the specification of a parametric form for F(x).

IX. APPENDIX B

If a quadratic function $y = \beta_1 + \beta_2 x_i + \beta_3 x_i^2$ $(i = 1, \ldots, k)$ is used to represent the log dose-response relationship in a bioassay, allowance must be made for the fact that such a response function must be monotonic; say non-decreasing, in the observed range (x_1 to x_k). In this appendix, a numerical example illustrating the fitting of a non-decreasing quadratic regression function by least squares is presented. Theoretical development of the general procedure for fitting non-decreasing polynomials was given by Lewish (23), in 1963.

The problem in this case is to estimate the parameters of $y = \beta_1 + \beta_2 x_i + \beta_3 x_i^2 + e_i$, $e_i \sim Nl(o, \sigma^2)$ subject to the linear constraints

$$\beta_2 + 2\beta_3 x_1 \geq 0$$
$$\beta_2 + 2\beta_3 x_k \geq 0$$

where the x_i's are known. These constraints insure that the first derivative is non-negative (and therefore that the quadratic function is non-decreasing) between x_1 and x_k.

Suppose the following hypothetical observations were made.

$x_1 = 0$	$y_1 = 2$
$x_2 = 1$	$y_2 = 7$
$x_3 = 2$	$y_3 = 9$
$x_4 = 3$	$y_4 = 10$

The unconstrained least square estimates in this case are

$\hat{\beta}_1 = 2.10$, $\hat{\beta}_2 = 5.60$, $\hat{\beta}_3 = -1.00$. The second constraint is violated in this case since $\hat{\beta}_2 + 2\hat{\beta}_3(3) = 5.60 - 6.00 = -0.4$. If both constraints had been satisfied, of course the unconstrained solution would be equivalent to the constrained solution.

The general method for determining the constrained solution, $\tilde{\beta}$, is first to make an orthogonal transformation from the β-space to γ-space, as described by Lewish, and then make an orthogonal projection of $\hat{\gamma}$ (unconstrained) to the parameter space, as defined by the linear constraints. Hartley (19) describes the construction of a triangular matrix A for making such a transformation, where $\gamma = A^{-1}\beta$.

The matrix A in this example is

$$A = \begin{bmatrix} \frac{1}{2} & \frac{-3}{2\sqrt{5}} & \frac{1}{2} \\ 0 & \frac{1}{\sqrt{5}} & \frac{-3}{2} \\ 0 & 0 & \frac{1}{2} \end{bmatrix}$$

The unconstrained $\hat{\gamma}$'s may be found by the relationship

$$\hat{\gamma} = A^{-1}\hat{\beta} = \begin{bmatrix} 14.000 \\ 5.814 \\ -2.000 \end{bmatrix}$$

where

$$A^{-1} = \begin{bmatrix} 2 & 3 & 7 \\ 0 & \sqrt{5} & \frac{15}{\sqrt{5}} \\ 0 & 0 & 2 \end{bmatrix}$$

The linear constraints in this case are

$$\beta_2 \geq 0$$

$$\beta_2 = 6\beta_3 \geq 0 \ .$$

By the transformation $A\gamma = \beta$, these constraints in the γ-space become

$$.447\ \gamma_2 - 1.500\ \gamma_3 \geq 0$$

$$.447\ \gamma_2 + 1.500\ \gamma_3 \geq 0 \ .$$

The parameter space is, therefore, bounded by the lines

$$\gamma_2 = 3.354\ \gamma_3$$

$$\gamma_2 = -\ 3.354\ \gamma_3$$

The point $\hat{\gamma}_2, \hat{\gamma}_3$ (5.814, - 2.000) lies outside the parameter space but close to the boundary line $\gamma_2 = -3.354\ \gamma_3$. The orthogonal projection from this point to the line will determine the constrained solution, $\tilde{\gamma}$. It can be seen that

$$\tilde{\gamma}_3 = \frac{\hat{\gamma}_2 - (3.354)(\hat{\gamma}_2)}{(3.354)^2+1}$$

$$= \frac{-2.000 - (3.354)(5.814)}{12.249}$$

$$= -1.76,$$

and $$\tilde{\gamma}_2 = 5.90.$$

Thus

$$\tilde{\gamma} = \begin{bmatrix} 14.00 \\ 5.90 \\ -1.76 \end{bmatrix}.$$

The transformation of $\tilde{\gamma}$ back into the β-space by the transformation $\tilde{\beta} = A\tilde{\gamma}$ yields

$$\tilde{\beta} = \begin{bmatrix} 2.16 \\ 5.28 \\ -0.88 \end{bmatrix}$$

which is the least square solution under the constraints. It can be seen that both constraints on the β's are now satisfied, since $\tilde{\beta}_2 \geq 0$ and $\tilde{\beta}_2 + 6\tilde{\beta}_3 = 0$.